COGNITIVE FUNCTION AND BREAST CANCER:

GENOMICS AND DISEASE CHARACTERISTICS

by

Theresa A. Koleck

B.S.N., University of Pittsburgh, 2011

Submitted to the Graduate Faculty of the

School of Nursing in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

SCHOOL OF NURSING

This dissertation was presented

by

Theresa A. Koleck

It was defended on

April 11, 2016

and approved by

Catherine M. Bender, PhD, RN, FAAN, Professor, School of Nursing

Susan M. Sereika, PhD, Professor, School of Nursing

Christopher M. Ryan, PhD, Professor Emeritus, Department of Psychiatry

Beth Z. Clark, MD, Assistant Professor, School of Medicine, Department of Pathology

Dissertation Advisor: Yvette P. Conley, PhD, Professor, School of Nursing

COGNITIVE FUNCTION AND BREAST CANCER:

GENOMICS AND DISEASE CHARACTERISTICS

Theresa A. Koleck, PhD, BSN, RN

University of Pittsburgh, 2016

Cognitive dysfunction is one of the most common and burdensome symptoms experienced by breast cancer survivors. This exploratory, ancillary study investigated the hypothesis that heterogeneity in the biology of breast cancers, including differences in clinicopathologic tumor features (CTFs) and host DNA variation in genes used clinically for breast cancer prognostication, may account for a proportion of variability in pretreatment (i.e., prior to initiation of systemic adjuvant therapy) cognitive performance among postmenopausal women diagnosed with early-stage breast cancer. The parent study, *Cognitive Impairment Related to Anastrozole Use in Women*, provided pretreatment cognitive function data, CTF information from surgical pathology reports of women with breast cancer prescribed to initiate anastrozole±chemotherapy (n=329) at a future time, and biospecimens for the women with breast cancer (n=138) and age- and education-matched healthy controls (n=82), who were identically assessed for cognitive function. We genotyped 131 single nucleotide polymorphisms (SNPs) representing 25 breast cancer-associated candidate genes (*AURKA, BAG1, BCL2, BIRC5, CCNB1, CD68, CENPA, CMC2, CTSL2, DIAPH3, ERBB2, ESR1, GRB7, GSTM1, MELK, MKI67, MMP11, MYBL2, NDC80, ORC6, PGR, RACGAP1, RFC4, RRM2,* and *SCUBE2*). Genetic risk/protection scores (GRSs) were calculated for each cognitive function composite to evaluate the collective effect of possession of multiple SNPs on cognitive performance. Multiple linear regression modeling was used to determine if CTFs, SNPs, and/or interactions between

CTFs and GRSs accounted for variability in cognitive performance. We found that CTFs related to cancer stage, tumor size, tumor focality, tumor location, histologic type and grade, hormone receptor and HER2 expression, cellular proliferation, as well as Oncotype DX® Breast Cancer Assay Recurrence Score® and Magee Equation recurrence score were individually significantly ($p<0.05$) associated with performance for one or more cognitive function composites. With the exception of *CMC2*, *MMP11*, and *RACGAP1*, significant ($p<0.05$) SNP main effect and/or SNP-by-prescribed treatment group interactions were observed individually between at least one cognitive function composite and one or more SNPs. Each GRS was significantly ($p<0.001$) associated with its respective cognitive function composite score. The findings from this dissertation study lay the foundation for a line of research to identify pathophysiologic mechanisms of and clinically relevant biomarkers for breast cancer-related cognitive dysfunction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Theresa would like to acknowledge her dissertation chair, committee members, family, peers, and the following support that made this dissertation study possible: Cognitive Function and Breast Cancer: Genomics and Disease Characteristics (F31NR014590), American Cancer Society Doctoral Degree Scholarship in Cancer Nursing (DSCN-14-076-01-SCN), Nightingale Awards of Pennsylvania PhD Degree Scholarship, Sigma Theta Tau International Eta Chapter Research Award, Ruth and Bill Fincke PhD Student Research Award, and Targeted Research and Academic Training Program for Nurses in Genomics (T32NR009759).

1.0 PROPOSAL INTRODUCTION AND SPECIFIC AIMS

The overall purpose of this dissertation study, *Cognitive Function and Breast Cancer: Genomics and Disease Characteristics*, is to gain a greater understanding of the biological foundations of cancer- and treatment-related cognitive decline in early-stage breast cancer survivors. Based on an extensive review of the literature and preliminary work conducted by the student and her mentoring team, it is hypothesized that differences in clinicopathologic tumor features (CTFs) and DNA variation in genes used to characterize the biology of breast cancer will be associated with changes in cognitive function in women with early-stage breast cancer. The specific aims of this dissertation study are threefold and depicted in the conceptual model (Figure 1):

Specific Aim 1: Investigate the relationship between CTFs of breast cancer and cognitive function in postmenopausal women with early-stage breast cancer.

Specific Aim 2: Explore DNA variation in genes used to clinically evaluate the biology of breast cancer for association with susceptibility to or protection from cognitive decline in postmenopausal women with early-stage breast cancer.

Specific Aim 3: Explore interactions between CTFs of breast cancer and DNA variation in genes used to clinically evaluate the biology of breast cancer on cognitive function.

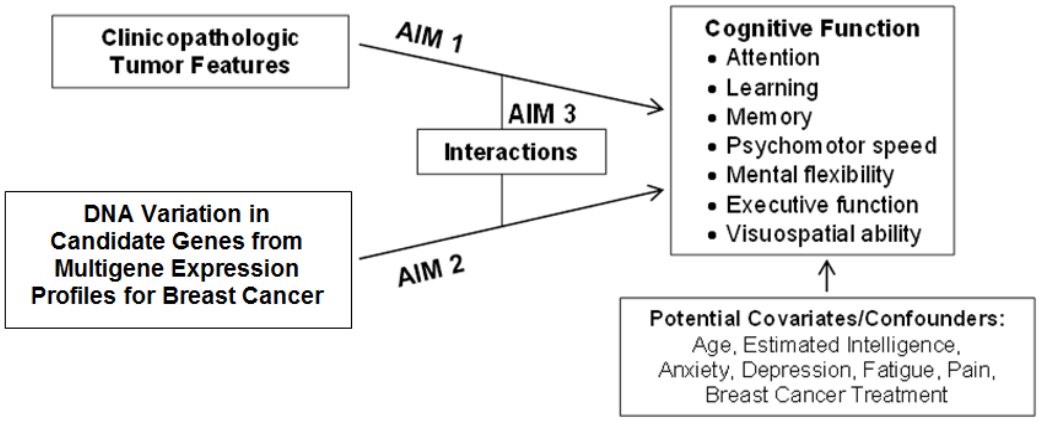


Figure 1: Conceptual Model

The proposed dissertation study is an ancillary study to a large, ongoing longitudinal study, *Cognitive Impairment Related to Anastrozole Use in Women (AIM) Study* (R01CA107408; PI: Dr. Catherine Bender). The AIM Study explores the impact of the adjuvant (i.e., therapy after primary surgery to reduce disease recurrence and overall mortality), anti-estrogen therapy, anastrozole, on changes in cognitive function longitudinally (prior to initiation of adjuvant therapy, every six months throughout the first two years of therapy, annually for the final three years of therapy and twelve months after the conclusion of therapy) in four cohorts of postmenopausal women: 1) women which breast cancer who receive chemotherapy plus anastrozole; 2) women with breast cancer who receive chemotherapy alone; 3) women with breast cancer who receive anastrozole alone; and 4) a healthy control group of women frequency-matched on age and years of education to the cohorts of women with breast cancer. Specifically, the AIM Study will provide data on demographics, certain CTFs, mood, fatigue, pain, breast cancer treatment, and cognitive function. The AIM Study will also provide banked DNA samples. DNA samples were obtained as part of a separate ancillary study (PIs: Drs. Bender and Yvette Conley) that allowed for the collection and biobanking of genetic specimens

from a subset of AIM Study participants. Drs. Bender and Conley have given the student permission to access the biobank and databank for her dissertation study. New genotype data will be generated using the biobanked DNA samples, and additional CTF data will be collected from the medical records of AIM Study participants for the proposed study. Please note that previous genetic investigations using the biobanked specimens have focused on DNA repair and oxidative stress genetic variants known to affect cognitive functioning. The proposed ancillary study adds and extends previous investigations through an innovative examination of genes used to characterize the biology of breast cancer.

1.1 SIGNIFICANCE AND INNOVATION

Breast cancer is the most prevalent form of cancer, excluding skin cancer, among women in the United States with an estimated 232,670 new cases of invasive breast cancer and 62,570 new cases of carcinoma in situ diagnosed in 2014 (American Cancer Society, 2014). Fortunately, advancements in detection and treatment have decreased breast cancer mortality rates and breast cancer survivors represent the largest group of cancer survivors in the United States at 2.8 million women (American Cancer Society, 2014). This increase in survival has transformed the care of breast cancer from that of a lethal diagnosis to a manageable, chronic disorder, accompanied by many burdensome cancer- and treatment-related symptoms. Changes in cognitive function or cognitive decline, defined as a decrease or loss in one or more of the domains of cognitive function including attention, learning, memory, psychomotor speed, mental flexibility, executive function, and visuospatial ability, is one of the most common and problematic phenomena experienced by breast cancer survivors. In fact, between 21 and 95% of

breast cancer survivors report some form of subjective cognitive complaints following adjuvant chemotherapy and/or anti-estrogen therapy (Downie, Mar Fan, Houédé-Tchen, Yi, & Tannock, 2006; Hurria et al., 2006; Jenkins et al., 2006; Mehnert et al., 2007; Schagen et al., 1999; Schilder et al., 2009; Shilling & Jenkins, 2007). These reported deficits have been objectively verified (Ahles et al., 2002; Bender et al., 2006, 2010; Brezden, Phillips, Abdolell, Bunston, & Tannock, 2000; Castellon et al., 2004; Fan et al., 2005; Jansen, Dodd, Miaskowski, Dowling, & Kramer, 2008; Jenkins et al., 2006; Jenkins, Shilling, Fallowfield, Howell, & Hutton, 2004; Lejbak, Vrbancic, & Crossley, 2010; Paganini-Hill & Clark, 2000; Quesnel, Savard, & Ivers, 2009; Schagen et al., 1999; Shilling, Jenkins, Fallowfield, & Howell, 2003; Tager et al., 2010; Tchen et al., 2003; van Dam et al., 1998; Von Ah et al., 2009; Wefel, Saleeba, Buzdar, & Meyers, 2010), and while the majority of breast cancer survivors do not exhibit profound cognitive impairments (Bender et al., 2006; Brezden et al., 2000; Castellon et al., 2004; Jim et al., 2009; Stewart, Bielajew, Collins, Parkinson, & Tomiak, 2006), even subtle changes in cognitive function can have a major impact on a survivor's quality of life, affecting relationships with family and friends, educational and career decisions, job performance, emotional state, ability to make informed treatment decisions, and adherence to cancer therapy (Bender et al., 2013; Boykoff, Moieni, & Subramanian, 2009; Munir, Burrows, Yarker, Kalawsky, & Bains, 2010; Myers, 2012; Tchen et al., 2003; Von Ah, Habermann, Carpenter, & Schneider, 2012).

More recently, studies that incorporate a pretreatment evaluation have found that compared to healthy controls or published normative data, women with breast cancer have poorer cognitive function prior to initiation of adjuvant chemotherapy and/or anti-estrogen therapy even after controlling for potential confounders, including age, estimated intelligence, depression, anxiety, and fatigue (Bender et al., 2006; Hermelink et al., 2007; Schilder et al.,

2010; Wefel et al., 2004, 2010), suggesting that inherent disease-related factors of breast cancer may contribute to the cognitive decline seen in women with breast cancer prior to adjuvant treatment (Ahles et al., 2011; Jim et al., 2009; Seigers et al., 2010; Wefel et al., 2004). Additionally, there is considerable variation in the incidence of cognitive decline among women, the severity of the cognitive decline, and the specific cognitive domains affected (Falleti, Sanfilippo, Maruff, Weih, & Phillips, 2005; Wefel & Schagen, 2012). These incongruities across findings of studies and between individual women lend themselves to and have the potential to be explained by genetic variation (Ahles & Saykin, 2007; Ahles et al., 2003; Small et al., 2011; Vardy, Wefel, Ahles, Tannock, & Schagen, 2008).

The proposed study aims to be an initial and integral step in establishing a relationship between traditional CTFs of breast cancer, DNA variability within genes used to clinically evaluate the biology of breast cancer, and cancer- and treatment-related cognitive decline. The proposed study will build upon the student's previous research experiences, as part of her mentoring team, investigating relationships between genes known to affect cognitive function and cognitive decline in women with breast cancer through exploration of a complementary, but novel dissertation hypothesis supported by the scientific literature: inherent, disease-related clinicopathologic and genetic features of breast cancer tumors can be used to personalize the prediction of cognitive decline in women with and receiving treatment for breast cancer. The proposed study, as well as the innovative line of investigation that this foundational study has the potential to generate, has significant implications for breast cancer survivors as (1) the results will add to our understanding of changes in cognitive function of women with and receiving treatment for breast cancer and (2) the knowledge generated has the potential to expand the utility of traditional CTFs and prognostic multigene expression profiles, indicate clinically

relevant biomarkers of susceptibility to or protection from cognitive decline in women with and receiving treatment for breast cancer, inform research into novel treatments for women experiencing cognitive decline, and provide a means for determining which women are most at risk for cognitive decline and would potentially benefit from earlier and/or more intensive interventions.

1.2 BACKGROUND

A comprehensive, systematic literature review (not limited by year of publication) was conducted of professional, published literature. Combinations of key words were entered into three electronic databases, PubMed, Ovid MEDLINE, and Google Scholar (Figure 2). All results produced by PubMed and Ovid MEDLINE and the first four pages of results on Google Scholar were reviewed.

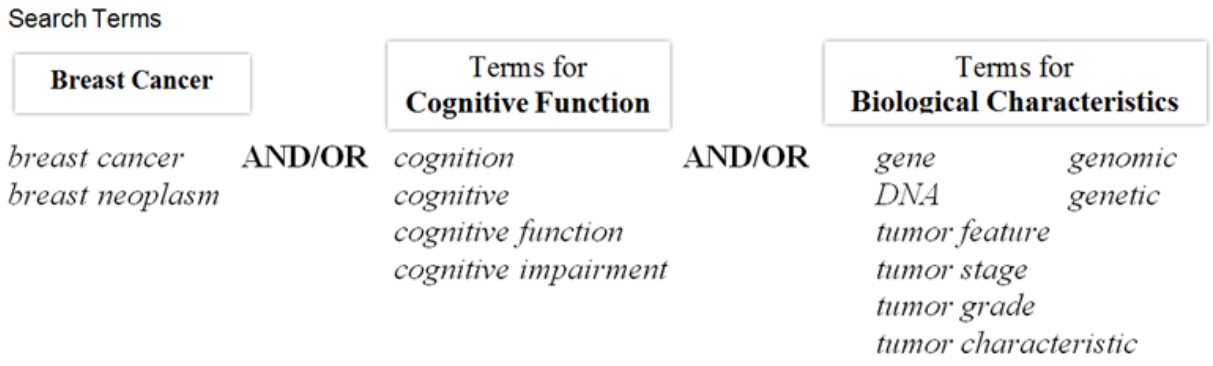


Figure 2: Literature Review Search Terms

1.2.1 Support for investigation of CTFs and cognitive decline

Ahles et al. found that at pretreatment, women with early-stage breast cancer (Stages 1-3) had significantly poorer reaction time than healthy controls; in contrast, subjects with noninvasive cancer (Stage 0) did not perform more poorly than healthy controls (2011). In a preliminary analysis of CTFs in a subset of available AIM Study data, it was found that positive estrogen receptor status, higher tumor grade, and high Oncotype DX® Breast Cancer Recurrence Score® predicted poorer pretreatment cognitive function in women with breast cancer. These findings all suggest that CTFs may impact cognitive function in women with early-stage breast cancer and warrant more extensive investigation.

1.2.2 Support for investigation of a genetic component of cognitive decline

While limited in number, studies have explored the relationship between DNA genetic variation and cognitive function in women with breast cancer and provide support for the hypothesis that DNA variation has the potential to influence cognitive function. Small et al. found that catechol-o-methyltransferase valine carriers treated with chemotherapy performed more poorly on tests of attention than healthy women who were also valine carriers (Small et al., 2011). Ahles et al. also found that genetic variation affects cognitive function in breast cancer survivors revealing that in a cohort of long-term breast cancer and lymphoma survivors treated with standard-dose chemotherapy, the *APOE* E4 allele is associated with poorer visual memory, spatial ability, and psychomotor functioning (Ahles et al., 2003).

1.3 PRELIMINARY STUDIES

To support and extend the previously reported evidence, the student and her mentoring team's own analysis of the relationship between cognitive function in postmenopausal women with breast cancer and *APOE* genotype revealed that pretreatment performance and/or changes in performance on tasks of executive function, attention, verbal learning and memory, and visual learning and memory were found to be influenced by *APOE* genotype and/or interactions between *APOE* genotype and breast cancer treatment (Koleck et al., 2014; Appendix F).

An analysis of 39 functional or tagging polymorphisms of select oxidative stress (*CAT*, *GPX1*, *SEPP1*, *SOD1*, and *SOD2*) and DNA repair (*ERCC2*, *ERCC3*, *ERCC5*, and *PARP1*) genes and pretreatment cognitive function was also conducted. Each cognitive function composite evaluated as part of the study was significantly associated with one or more oxidative stress and DNA repair gene polymorphisms through either single nucleotide polymorphism (SNP) main effects and/or SNP-by-prescribed breast cancer treatment group interactions, further suggesting that genetic variation has the potential to influence cognitive function (Koleck et al., 2016; Appendix J).

1.3.1 Milestones

The following table (Table 1) lists the milestones that have been achieved since entrance into the BSN-to-PhD program at the University of Pittsburgh School of Nursing in May 2011. All milestones support the scientific merit of the proposed dissertation study.

Table 1: BSN-to-PhD Program Milestones

Milestone	Date
Targeted Research and Academic Training Program for Nurses in Genomics (T32NR009759) fellowship appointment	May 2011
University of Pittsburgh School of Nursing Bessie Li Sze Memorial Scholarship	Aug 2012 Aug 2013
University of Pittsburgh Institutional Review Board Approval for *Cognitive Function and Breast Cancer: Genomics and Disease Characteristics* (expedited review, PRO13040672; Appendix C)	June 2013
Preliminary Examination	July 2013
Cognitive Function and Breast Cancer: Genomics and Disease Characteristics (F31NR014590) funded by National Institute of Nursing Research	Dec 2013
Sigma Theta Tau International Eta Chapter Research Award	Mar 2014
Nightingale Awards of Pennsylvania PhD Degree Scholarship	Apr 20214
American Cancer Society Doctoral Degree Scholarship in Cancer Nursing (DSCN-14-076-01-SCN)	July 2014
University of Pittsburgh School of Nursing Ruth and Bill Fincke PhD Student Research Award	July 2014
Comprehensive Examination and Overview	Dec 2014

1.4 DESIGN AND METHODOLOGY

The proposed ancillary study will utilize regression modeling (Aim 1), candidate gene association (Aim 2), and moderation analyses (i.e., statistical interactions) (Aim 3) to explore changes in cognitive function in cohorts of women with breast cancer prior to and over time following the initiation of adjuvant therapy. The anticipated timeline for completion of study aims is provided below (Table 2).

Table 2: Study Timeline

	Year 1 5/11 – 4/12	**Year 2** 5/12 – 4/13	**Year 3** 5/13 – 4/14	**Year 4** 5/14 – 4/15	**Year 5** 5/15-4/16
Coursework					
Summer Genetics Institute *June 2012*					
F31 Submitted *Dec 2012*					
IRB Approval *June 2013*					
Preliminary Exam Completed *July 2013*					
F31 Funding *Dec 2013 – Apr 2016*					
Genotype Data Collection *at the University of Pittsburgh School of Nursing Molecular Genomics Laboratory*					
Admission to Candidacy *Dec 2014*					
Abstraction of Clinicopathologic Tumor Feature Data *from medical records at University of Pittsburgh Medical Center (UPMC) facilities*					
Data Analysis and Interpretation					
Preparation of Research Presentations and Manuscript Development					
Dissertation Defense *April 2016*					

1.4.1 Setting and sample

As the proposed study is an ancillary study to the AIM Study, the target population and eligibility criteria are consistent with those of the AIM Study. Women with breast cancer were recruited from the Comprehensive Breast Cancer Program of the University of Pittsburgh Cancer Institute. Healthy women were recruited using a variety of approaches including referral from

women in the breast cancer cohorts, advertisements, and random digit dialing. Inclusion criteria consist of being female, postmenopausal (defined as amenorrhea persisting for an entire year, oophorectomy, or hysterectomy and age greater than 51 years, the average age of menopause in the United States), maximum age of 75 years, able to speak and read English, and completion of a minimum of 8 years of education. Exclusion criteria for all participants consist of self-report of hospitalization for psychiatric illness within the last 2 years and having a prior diagnosis of neurologic illness. Additional inclusion criteria for participants with breast cancer are being diagnosed with early-stage (1, 2, or 3a) breast cancer based on the TNM Classification of Malignant Tumors system (Edge et al., 2010) with confirmation by each participant's medical oncologist, and eligible to receive either chemotherapy plus anastrozole, chemotherapy alone, or anastrozole alone. Further exclusion criteria for participants with breast cancer are clinical evidence of distant metastases or a prior diagnosis of cancer. Additional inclusion criteria for the control group include no current or history of any cancer and not currently taking (for at least 3 months) any form of hormone replacement therapy.

1.4.2 Definition of cognitive decline

Cognitive decline will be defined as poorer performance on cognitive measures within participants over time or poorer performance on cognitive measures in women with breast cancer compared to healthy controls. Cognitive function was evaluated at each time point using a comprehensive neuropsychological test battery (Table 3). In order to mitigate the influence of practice effects, cognitive tests with alternate, equivalent versions were administered at follow-up testing. Scores from the control group also allow for statistical control of potentially confounding practice effects. Please note that data reduction methods are conducted as part of

the AIM Study to reduce the dimensionality of the cognitive function data. Eight cognitive function composites were identified and composite z-scores will be utilized as the dependent/outcome variables in this study: attention, concentration, executive function, mental flexibility, psychomotor speed, verbal memory, visual memory, and visual working memory.

Table 3: AIM Study Neuropsychological Test Battery

Cognitive Domain	Neuropsychological Test	Reliability and Validity References
Attention	*Digit Vigilance Test* (Lafayette Clinical Insturments Company, 1989)	Bender et al., 2000; Ryan, Williams, Orchard, & Finegold, 1992; Ryan & Williams, 1993
	CANTAB Rapid Visual Information Processing Test (Owen, Sahakian, Semple, Polkey, & Robbins, 1995)	Sahakian & Coull, 1993
Learning and Memory	*CANTAB Paired Associates Learning Test* (Owen et al., 1995)	Lowe & Rabbitt, 1998; Owen et al., 1995; Robbins et al., 1997
	Rivermead Behavioral Memory Test (Wilson, Cockburn, Baddeley, & Hiorns, 1989)	Cockburn & Smith, 1993; Jenkins et al., 2004; Wilson et al., 1989
	Rey Auditory Verbal Learning Test (Rey, 1941)	Powell, Cripe, & Dodrill, 1991
	Rey Complex Figure Test (Osterrieth, 1944)	Berry, Allen, & Schmitt, 1991
	CANTAB Spatial Working Memory Test (Owen et al., 1995; Robbins et al., 1997)	Lowe & Rabbitt, 1998; Owen et al., 1995
Psychomotor Speed	*Grooved Pegboard Test* (Lafayette Clinical Insturments Company, 1989)	Matthews, Cleeland, & Hopper, 1970; Wieneke & Dienst, 1995
	Digit Symbol Substitution Test (Wechsler, 1981)	Jenkins et al., 2004; Snow, Tierney, Zorzitto, Fisher, & Reid, 1989; Youngjohn, Larrabee, & Crook, 1992
Mental Flexibility	*Trail Making Test-B* (Reitan & Wolfson, 1985)	Bornstein, 1985; Moertel, Reitmeier, Bolton, & Schorter, 1964; Reitan, 1958; Wieneke & Dienst, 1995
	D-KEFS Color-Word Interference Test (Delis, Kaplan, & Kramer, 2001)	Delis et al., 2001
Executive Function	*CANTAB Stockings of Cambridge Test* (Owen et al., 1995; Robbins et al., 1997)	Capuron, Ravaud, & Dantzer, 2001; Lowe & Rabbitt, 1998; Owen et al., 1995
	D-KEFS Verbal Fluency Test (Delis et al., 2001)	Delis et al., 2001
Visuospatial Ability	*Rey Complex Figure Test* (Osterrieth, 1944)	Berry et al., 1991; Wieneke & Dienst, 1995

Note. CANTAB=Cambridge Neuropsychological Test Automated Battery; D-KEFS=Delis-Kaplan Executive Function System

1.4.3 Clinicopathologic tumor features (CTFs)

CTFs, independent variables of interest, are used to characterize breast cancer tumor pathology. CTF data not originally collected as part of the AIM Study will be obtained from pathology reports. A number of CTFs will be evaluated and are detailed in Table 4.

Table 4: CTF Description

Feature	Classification/Evaluation Method	Scale
Overall TNM Stage	Stage 1, 2a, 2b, or 3a (Edge et al., 2010)	Nominal
Tumor Stage	Stage T1a, T1b, T1c, T2, T3 (Edge et al., 2010)	Nominal
Lymph Node Involvement	Positive or negative Number positive nodes	Nominal Ratio
Tumor site	Laterality: left or right breast Clock position/octant location: upper outer, upper inner, lower outer, lower inner, upper middle, lower middle, inner middle, outer middle	Nominal Nominal
Tumor focality	Single or multiple	Nominal
Primary tumor size	Measured to nearest millimeter	Ratio
Aggregate tumor size	Measured to nearest millimeter	Ratio
Histologic type	Invasive ductal, invasive lobular, both (Tavassoli & Devilee, 2003)	Nominal
Histological grade	Glandular/tubular differentiation score (1-3) Nuclear pleomorphism score (1-3) Mitotic activity/count score (1-3) Nottingham Score (3-9) Nottingham Grade (Elston & Ellis, 1991): Grade 1 (low), 2 (intermediate), or 3 (high)	Ordinal Ordinal Ordinal Ordinal Nominal
LV invasion	Presence or absence	Nominal
Estrogen and progesterone receptor	Positive or negative H-score quantitation (0-300) Oncotype DX quantitative gene scores (Genomic Health Inc., 2016) ER: 0-12.5+; Negative <6.5, Positive ≥6.5 PR: 0-10+; Negative <5.5, Positive ≥5.5	Nominal Ratio Ratio
HER2/neu	Positive or negative IHC classification score (Wolff et al., 2007): 0, 1+ (Negative), 2+ (Equivocal), or 3+ (Positive) Oncotype DX quantitative gene scores (Genomic Health Inc., 2016) 0-13+; Negative <10.7, Equivocal 10.7-11.5, Positive ≥11.5	Nominal Ordinal Ratio
Ki67	Classification: low, moderate, high, or very high % Index (0-100)	Nominal Ratio
Oncotype DX Recurrence Score®	Prognostic gene expression algorithm (0-100) (Paik et al., 2004, 2006)	Interval

Note. HER2=Human epidermal growth factor receptor 2; IHC=Immunohistochemistry; LV=Lymphovascular; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay.

1.4.4 Justification of candidate gene selection

One common clinical tool used to evaluate the underlying biology of breast cancer cells is the prognostic multigene expression profile for breast cancer. Several multigene expression profiles have been developed, including the 11-gene expression signature, marketed as the Breast Cancer Index℠ by bioTheranostics, Inc., San Diego, CA (Jerevall et al., 2011); the 14-gene prognostic expression signature described in Tutt et al. (2008); the 21-gene breast cancer assay, marketed as the Oncotype DX® Breast Cancer Assay by Genomic Health, Inc., Redwood City, CA (Paik et al., 2004, 2006); the 50-gene breast cancer prognostic gene signature assay, marketed as the Prosigna® Breast Cancer Prognostic Gene Signature Assay by NanoString Technologies, Inc., Seattle, WA (Dowset et al., 2013; NanoString Technologies Inc., 2013), based on the PAM50 Breast Cancer Intrinsic Classifier (Parker et al., 2009); and the 70-gene breast cancer recurrence assay, marketed as the MammaPrint® 70 Gene Breast Cancer Recurrence Assay by Agendia, Irvine, CA (Buyse et al., 2006; Veer et al., 2002). Briefly, multigene expression profiles for breast cancer enhance the knowledge received from traditional CTFs and utilize predictive algorithms to clinically evaluate the underlying biology of the cancer cells, individualizing treatment through estimation of adjuvant therapy benefit and distant cancer recurrence or metastasis risk; thus, each profile is comprised of genes that play an important role in breast cancer development and progression, and, consequently, represent ideal candidates for a genetic association study exploring our hypotheses. Characteristics of the previously listed profiles and the prioritization of candidate genes included in these profiles are discussed in detail in the article “Identification and prioritization of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level” (Koleck & Conley, 2016; Appendix H).

1.4.5 Candidate genes and polymorphisms selected for investigation

As described in the article "Identification and prioritization of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level," genes utilized in two or more multigene expression profiles for breast cancer were prioritized for investigation (Koleck & Conley, 2016; Appendix H). The remaining genes used in the 21-gene breast cancer assay predictive algorithm were also included, as this profile is utilized by AIM Study recruitment sites. In total, 25 candidate genes were selected: *AURKA, BAG1, BCL2, BIRC5, CCNB1, CD68, CENPA, CMC2, CTSL2, DIAPH3, ERBB2, ESR1, GRB7, GSTM1, MELK, MKI67, MMP11, MYBL2, NDC80, ORC6, PGR, RACGAP1, RFC4, RRM2,* and *SCUBE2*.

Single nucleotide polymorphisms (SNPs) within these genes will be analyzed as independent variables. Please refer to Table 6 in Appendix A for a list of proposed SNPs for each candidate gene. Functional polymorphisms were first selected from the literature. When a functional polymorphism was not identified and/or did not fully represent the gene of interest, tagging SNPs were selected using the Phase III HapMap database. Because the profiles rely upon gene expression data, evaluation of DNA variability was extended ±2,500bps beyond the gene to capture the UTR-5' and UTR-3' regulatory regions. Criteria for selecting tagging SNPs included: R^2 of ≥ 0.8, minor allele frequency of ≥ 0.20 (except where otherwise noted), and selected for Caucasian ancestry, which represents the majority of AIM Study participants.

1.4.6 Covariates and confounders

Demographic information, including age, years of education, and estimated verbal intelligence (National Adult Reading Test-Revised) were collected at the initial study time point (Nelson, 1982). Treatment information and menopausal status was verified via the participant's medical record. Depression, Beck Depression Inventory-II (Beck, Steer, & Brown, 1996); anxiety, POMS (Profile of Mood States) Tension-Anxiety subscale (McNair, Lorr, & Droppleman, 1992); fatigue, POMS Fatigue-Inertia subscale (McNair et al., 1992); and pain, Brief Pain Inventory (Cleeland, 1989) were also evaluated at each time point.

1.4.7 Genotype data collection and quality checks

The choice of genotyping platform was determined based on the number of participants and the number of SNPs to be evaluated. The Sequenom® iPLEX® MassARRAY platform will be used for genotype data collection. This approach allows for high throughput genotyping using a multiplex approach. This platform is available through the University of Pittsburgh Clinical and Translational Science Institute supported Genomics and Proteomics Core Laboratory. Blind duplicates are included within each assay to control for plate to plate variability. SNPs not amenable to multiplexing or with less than a 90% call rate will be (re)genotyped with a different platform, specifically ABI TaqMan® allelic discrimination, or eliminated from the analyses. Each SNP will be tested for Hardy-Weinberg Equilibrium in an effort to detect potential genotyping errors.

1.4.8 Sample size justification

Three-hundred and fifty-one women with early-stage breast cancer and 136 healthy controls have complete baseline/pretreatment assessments in the AIM Study. Of these participants, 242 have banked DNA and demographic and cognitive function data for one or more assessment time points. Due to depreciation in sample size, the current study will focus on the first three time points. Current sample breakdown for T0, T1, and T2 are presented in Table 5. All CTFs (Aim 1), with the exception of Oncotype DX® Breast Cancer Assay data (only available for eligible women who completed testing), are evaluated as part of a standard pathology report and will be analyzed in all study participants with breast cancer. Genetic variation in candidate genes used to clinically evaluate the biology of breast cancer (Aim 2) will be explored in all women with a biobanked genetic sample including healthy controls. Interactions between CTFs and DNA variation (Aim 3) will be analyzed in an exploratory manner in women with breast cancer who also have a biobanked genetic sample. Based on our preliminary analyses and available sample, the proposed study would have .80 power to detect small effect sizes (R^2) of 0.022 and 0.020 for individual CTFs without adjustment and with adjustment of a set of 8 covariates explaining moderate variability in the outcome variable (R^2=0.09), respectively, for Aim 1 using an F-test at a significance level of 0.05 for two-sided hypothesis testing. When modeling the change in cognitive function using linear mixed-effects regression with three time points, effects for individual CTFs as small as f=0.16 can be detected with 0.80 power at a two-tailed significance level of 0.05. For Aim 2, a sample size of 242 achieves 0.80 power to detect small effect sizes (R^2) of 0.031 and 0.025 for individual candidate genes without adjustment and with adjustment of a set of 8 covariates explaining variability in the outcome variable of R^2=0.20, respectively, using an F-test with a significance level of 0.05. Interactions between CTFs and DNA variation

will be analyzed in an exploratory manner in women with breast cancer who also have a biobanked genetic sample.

Table 5: AIM Study Sample Size (Genetic Sample Size) by Time Point

	T0	T1	T2
Chemotherapy plus Anastrozole	125 (55)	101 (56)	68 (49)
Chemotherapy Alone	28 (15)	26 (15)	24 (15)
Anastrozole Alone	198 (83)	139 (79)	94 (66)
Healthy Controls	136 (83)	111 (77)	95 (64)
Total	487 (236)	377 (227)	281 (194)

Note. All women with breast cancer completed a pretreatment (T0) cognitive assessment after primary surgery but prior to initiation of systemic adjuvant therapy. Women with breast cancer prescribed chemotherapy plus anastrozole completed follow-up cognitive assessments after chemotherapy but prior to initiation of anastrozole (T1) and six months after anastrozole initiation (T2). Women with breast cancer prescribed chemotherapy alone completed follow-up assessments after chemotherapy (T1) and six months after completion of chemotherapy (T2). Women with breast cancer prescribed anastrozole alone completed follow-up cognitive assessments six (T1) and twelve (T2) months after initiation of anastrozole. Healthy controls were assessed at comparable time points, i.e., baseline (T0) and six (T1) and twelve (T2) months after baseline.

1.4.9 Data analysis

Results of interest for all aims are the estimated regression coefficients and 95% confidence intervals, and the tests of significance of these coefficients at a two-tailed significance level of 0.05. Estimates of the proportion of variance explained in cognitive function composites will also be obtained by each CTF and allele and/or genotype.

1.4.9.1 Descriptive statistics Standard descriptive statistics will be computed for all independent, dependent, and potentially confounding/covariate variables based on the variable's measurement level and observed distribution. For nominal data, simple frequencies, percentages, ranges and modes will be examined. Ordinal data descriptive statistics will further include semi- and inter-quartile ranges and medians. For continuous type (interval or ratio scaled) data, frequencies, variances, means, medians, and ranges, including standard deviations

and quartiles, will be described. Variable distributions will be examined using univariate or bivariate frequency distributions, such as cross-tabulation contingency tables, and appropriate graphical displays, including histograms, boxplots, and scatterplots. Differences in baseline confounder/covariate data by study cohort and/or genotype and/or CTF (with continuous CTFs being meaningful categorized) will be evaluated using factorial ANOVA to determine statistical and clinical significance of mean group differences. The comparability of baseline confounder/covariate data and baseline cognitive ability between participants included in the proposed study analysis and remaining participants from the AIM Study will be assessed using two-sample t-tests to evaluate equality of means.

1.4.9.2 Data screening procedures Data will be screened for anomalies prior to the analyses for study aims. This preliminary analysis will be used to assess 1) accuracy of inputted data, 2) potential outliers and influential points, 3) the amount and pattern of missing data, and 4) potential violation of assumptions necessary for the planned analyses.

Screening for accuracy will be accomplished through examination of descriptive statistics and graphical representations of variables. For continuous variables, minimum values, maximum values, means, and standard deviations will be examined for plausibility. For discrete variables, data will be assessed for out-of-range category values and inaccurately programmed missing value codes. Out-of-range values will be checked for accuracy, amended, and retained in the analysis if possible.

Missingness of data will first be examined by participant; if a participant is contributing very little useable information, she will be dropped from the analysis. Missing values for all variables at each time point will then be described using frequencies and percentages. Reasons for missing data (e.g., laboratory error, participant refusal, etc.) will be noted. A missing values

analysis will be undertaken to explore patterns of missingness and potential violations in assuming that data are missing completely at random (MCAR). Separate variance t-tests will be used to investigate possible systemic missingness between the missingness for one variable and the observed values of any other variable. Little's MCAR test will be used to determine if data are missing completely at random. If only a few cases (i.e., <5%) have missing data and appear to be a random subset of the entire sample, simple deletion of missing cases will be used to handle missingness. Alternatively, if data are missing from a large number of cases or missing values are not randomly distributed, an imputation method will be applied to estimate missing data. Imputation methods that will be considered include expectation maximization and multiple imputation.

Outliers not due to incorrect data entry or having incorrectly specified missing value indicators will also be investigated. Univariate and multivariate outliers for both continuous and discrete variables will be examined. For discrete variables, uneven category splits will be identified using frequency distributions. Because regression analyses are planned, ungrouped data will be used for identification of univariate outliers in continuous variables. Histograms and boxplots will be used to identify points far removed from the distribution. Z-scores will be calculated and cases with extreme values (i.e., <-3.29 or >3.29) will be noted as potential outliers. The presence of multivariate outliers will be assessed using bivariate scatterplots and Mahalanobis distances. Outliers and influential points will also be evaluated for each regression model generated. Outliers in Y will be assessed using jackknifed (deleted studentized) residuals, and outliers in X will be assessed using leverage statistics; generated values will be depicted graphically (i.e., residual/leverage value boxplots and residual by leverage plots) to elucidate potential outliers. To explore how an observation may impact predicted values and individual

regression coefficients, the influence of the i-th observation on predicted values (DFFITS) and individual regression coefficients (DFBETAS) will be calculated. In order to determine if the i-th observation exerts undue influence on a set of coefficients, Cook's distance will be calculated. Additionally, covariance ratios (COVARATIO) will be generated to determine if the i-th observation improves or worsens the estimation ability of the model.

To evaluate the robustness of findings, regression models excluding points determined to be influential, as well as a robust regression models utilizing Huber and biweight iterations, will be generated. Models eliminating potentially influential multivariate-outlier cases and/or diminishing the weight of potentially influential univariate-outlier cases will be created as part of the sensitivity analysis as well.

Underlying assumptions for each regression model will be assessed. To evaluate independence among residuals, jackknifed residuals will be plotted against fitted values. A Durbin-Watson test will also be conducted to detect potential autocorrelation. A histogram of jackknifed residuals will be created to evaluate the normality of residuals. Measures of skewness and kurtosis of jackknifed residuals will also be generated; values deviating from zero will be further investigated. Added variable plots and component-plus-residuals plots will be generated to assess linearity. Homoscedasticity will be evaluated using the Breusch-Pagan/Cook-Weisberg test for heteroscedasticity. Variance inflation factors and conditioning indices and variance decomposition proportions will be used to assess for multicollinearity.

1.4.9.3 Analysis for Specific Aim 1 Multiple linear regression modeling will be performed to evaluate the effect of each CTF on all cognitive function composites for each time point (i.e., T0, T1, and T2). Models will initially include single CTFs to yield unadjusted estimators. Subsequently, in order to obtain adjusted estimators (i.e., minimally confounded estimates of

effect), confounders/covariates will be considered for inclusion in each model in a hierarchical fashion. Age, estimated verbal intelligence, and study cohort will be included in regression models as potential time-invariant (i.e., fixed) confounders/covariates. Depression, anxiety, fatigue, and pain will be included as time-dependent confounders/covariates. Due to the known influence of age and estimated verbal intelligence on cognitive function, these variables will be incorporated in the first block of the hierarchical regression analysis and retained in all models regardless of statistical significance. The second block will include depression, anxiety, fatigue, and pain. Decisions on whether to retain or eliminate this set of predictors will be based on model R^2 change and influence on predicted regression coefficients. The third and fourth blocks will include the treatment regimen and a CTF, respectively, the predictors of interest. Because we are interested in how the effect of each CTF is potentially modified by treatment regimen, interactions between CTFs and study cohort will be initially examined in the multivariable model in an additional block, block 5. If no significant interactions are observed, the main effects model will be utilized. Similarly, linear mixed-effects regression models will be used to evaluate the effect of CTFs on each cognitive function factor over time.

1.4.9.4 Analysis for Specific Aim 2 Multiple linear regression modeling will also be used to examine the relationship between each cognitive function factor and the presence (i.e., heterozygous and homozygous minor variant) or absence (i.e., homozygous wild type) of a candidate gene allele at T0, T1, and T2. If the allelic distribution allows, relationships between each cognitive function factor and participant genotype will be examined to elucidate allele dosage effects. Models will be fitted considering each candidate gene's SNPs separately to yield unadjusted regression coefficients. Using the same hierarchical regression strategy as proposed for analysis of Aim 1, models will be expanded to include potential confounders/covariates to

yield adjusted regression coefficients. Linear-mixed effects regression modeling will be used to model the change in individual cognitive function factors over time as a function of the presence of an allele and/or genotype. In addition, genetic risk/protection scores (GRSs) will be calculated to evaluate the collective effect of multiple candidate gene variants on cognitive function in women with breast cancer. Both a simple count and weighted method will be employed. For the simple count method, values of 1 will be assigned to SNP variants associated with poorer cognitive function factor scores and values of -1 will be assigned to SNP variants associated with improved cognitive function factor scores. Assuming an additive model of SNP effect, values will be summed across SNPs to produce a GRS for each participant. The weighted method will utilize regression coefficients from univariate SNP models to assign greater risk to SNPs with stronger associations. GRS will be calculated for each candidate gene and overall across all candidate genes. GRS will be evaluated as predictors in the multiple linear regression models and linear mixed-effects regression models as the presence of a candidate gene allele was explained previously.

1.4.9.5 Analysis for Specific Aim 3 Multiple linear regression models and linear-mixed effects regression models will be expanded to include the main effects for both individual CTFs and the presence of a particular allele and/or genotype and their interaction(s). Interaction terms will be created as the products of individual CTFs and the presence of a particular allele and/or genotype.

1.5 POTENTIAL LIMITATIONS AND ALTERNATIVE APPROACHES

While the sample sizes are appreciable, the student acknowledges that this study aims to explore many variables and recognizes that it may be underpowered for some aspects of the study. Likewise, while the reported minimum detectable R^2s are comparable to the genetic effects found in our previous and ongoing investigations of cognitive function in women with breast cancer, research has shown that variation in individual candidate genes typically accounts for a small amount (often 1% or less) of the variance in any given observed phenotype, with the additive combination of multiple genetic polymorphisms (and/or multiple CTFs) explaining a clinically meaningful proportion of the variability. Therefore, results from this study will need to be interpreted with a degree of caution. Findings from this study will serve as the foundation for future investigations.

In the case that genotype data cannot be collected for a particular SNP, another highly informative SNP in linkage disequilibrium with the original polymorphism will be selected. In addition, if a priority SNP is not amenable to multiplexing with the Sequenom® iPLEX® MassARRAY platform and does not have an available TaqMan® allelic discrimination assay, alternative approaches including custom assay and restriction fragment length polymorphism design will be employed in an attempt to obtain genotype information.

1.6 HAZARDOUS MATERIAL AND PROCEDURES

While working in the laboratory, the student will be exposed to bodily fluids and chemicals. Personal protective equipment, including gloves and a laboratory coat, will be worn and safety

protocols will be followed. The student has received blood borne pathogen training and chemical hygiene training. All experiments will be conducted in an appropriately equipped laboratory.

1.7 RESEARCH PARTICIPANT RISK AND PROTECTIONS

This dissertation study was originally submitted, reviewed, and approved by the University of Pittsburgh Institutional Review Board (IRB) in June 2013 (PRO13040672; Appendix C).

1.7.1 Human subjects involvement, characteristics, and design

The proposed study will use biobanked, de-identified genetic samples from a genetic ancillary study to a larger, ongoing parent study (AIM Study) that explores cognitive decline in women with and receiving treatment for breast cancer. The ancillary study developed a biobank of genomic material from these women and has explored variability in candidate genes involved in DNA repair and protection from oxidative damage. Participants are still being actively recruited for the AIM Study and genetic samples continue to be collected and biobanked. Women with breast cancer are recruited from the Comprehensive Breast Care Program of the University of Pittsburgh Cancer Institute. Control group candidates are accrued by utilizing the services of the University Center for Social and Urban Research and by referral from women in the breast cancer groups.

1.7.2 Sources of materials

Genetic samples are/were collected from participants in the AIM Study in the following manner: Participants that previously completed cognitive function data collection for the AIM Study were re-contacted for the purpose of obtaining a biological sample for DNA extraction and biobanking. Only individuals who agreed to being re-contacted were re-contacted for the ancillary study. Participants that are currently undergoing data collection for the AIM Study are asked about their interest in participating in the ancillary genetics study during one of their data/sample collection visits. If they indicated interest in the genetic portion of the study, they were fully informed about the study prior to signing an informed consent document. It is important to note that for candidate gene analysis of DNA, it does not matter at what time point the genetic sample collection occurs; therefore, sample collection can occur at any time point. A sample of 3cc of whole blood or 2cc of saliva is collected from each participant who agrees to participate in the genetic portion of the study. Samples are processed and DNA is extracted and biobanked. The stock samples are placed in a -80°C freezer for long-term storage. The stock samples will be aliquoted and diluted for genotyping in this dissertation study.

1.7.3 Potential risks

As the proposed study involves the use of identifiable medical information, potential risks to study participants includes the risk of breach of confidentiality of data and the risk of anonymity of participants. Actions, as detailed below, are taken to keep any information obtained as confidential/anonymous as possible.

1.7.4 Protection against risk

A signed addendum informed consent for the collection and analysis of genetic samples was obtained from participants prior to sample collection. The use of the sample for investigation of genetic aspects of cognitive function related to breast cancer is explained to participants prior to obtaining their informed consent. Within the consent, participants agree to have their genetic material available for analysis of any gene(s) that may be involved in cognitive function within the context of breast cancer. The University of Pittsburgh IRB has approved the consent form and protocols used for recruitment and specimen collection. These documents are reviewed on an annual basis. All participants are assigned a unique code number under which all data are stored. Security of data is upheld through the use of password protection and restricted access to users. Consent forms and a list of the match between participant names and code names are retained in a discrete locked file cabinet in Dr. Catherine Bender's office. All staff and students who interact with study participants and/or who have access to participant identifiers are required to sign a confidentiality agreement and to complete online education modules sponsored by the Research Conduct and Compliance Office of the University of Pittsburgh prior to contact with any participants or access to medical records and/or data. Risk to the confidentiality of the genetic information generated by the proposed study is a very minimal risk given that all of the samples arrive in the laboratory already de-identified and all databases related to the genetic material will be de-identified. Each sample and piece of genetic data will be associated only with a unique code. Every precaution will be taken to minimize exposure of the data to persons outside the project by using passwords for all computer files and keeping all hard copies of data within the genomics laboratory, which is locked and accessible only to authorized personnel. Data from this project will be reported in aggregate only.

1.7.5 Potential benefits

Research participants will likely receive no direct benefit from taking part in the proposed research study. Nevertheless, the proposed dissertation study has the potential to improve cognitive function in women with and receiving treatment for breast cancer by indicating clinically relevant biomarkers of susceptibility to or protection from cognitive decline in women with and receiving treatment for breast cancer, inform research into novel treatments for women experiencing cognitive decline, and provide a means for determining which women are most likely at risk for cognitive decline and would potentially benefit from earlier and/or more intensive interventions. Considering the minimal risk to research participants, the proposed study is an advantageous undertaking.

2.0 SUMMARY OF STUDY

The purpose of this dissertation research was to: (1) investigate the relationship between clinicopathologic tumor features (CTFs) of breast cancer and cognitive function in postmenopausal women with early-stage breast cancer; (2) explore DNA variation in genes used to clinically evaluate the biology of breast cancer for associations with susceptibility to or protection from cognitive decline in postmenopausal women with early-stage breast cancer; and (3) explore interactions between CTFs of breast cancer and DNA variation in genes used to clinically evaluate the biology of breast cancer on cognitive function. Four articles, not including articles generated from the three aims noted above, that were written during the course of PhD training are provided in Appendices D, F, H, and J. The first article, published in the *Annual Review of Nursing Research* and entitled "Molecular genomic research designs" discusses key considerations for designing studies with a molecular genetic or genomic focus (Appendix D). The data-based article "Apolipoprotein E genotype and cognitive function in postmenopausal women with early-stage breast cancer," published in *Oncology Nursing Forum*, complements the hypothesis of this dissertation research and examines the role of *APOE* genotype in cognitive function of postmenopausal women with early-stage breast cancer prior to initiation of adjuvant therapy and over time with treatment (Appendix F). A second data-based manuscript, which also complements this dissertation research by examining relationships between polymorphisms in oxidative stress and DNA repair genes and pre-adjuvant therapy

cognitive function in postmenopausal women diagnosed with early-stage breast cancer, was published in *SpringerPlus* (Koleck et al., 2016; Appendix J). The forth manuscript, entitled "Identification and prioritization of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level," was published in *Breast Cancer: Targets and Therapy*. This article provides a detailed background to support the hypothesis underlying this dissertation work and discusses the novel approach that was employed to select candidate genes to test the hypothesis (Appendix H). The results for Specific Aims 1, 2, and 3 are presented in the data-based manuscript, "The impact of variation in clinicopathologic tumor features and breast cancer-related genetic polymorphisms on pretreatment cognitive function in women with breast cancer: an exploratory analysis," included in this document.

2.1 PRELIMINARY WORK ON DISSERTATION STUDY

Prior to the comprehensive examination and overview, significant progress was made in the collection of genotype data related to Aim 2. First, the student finalized single nucleotide polymorphism (SNP) selection, categorizing and ranking SNPs from highest to lowest priority within each gene in the following manner: 1) functional consequence, 2) tagging in a regulatory region, and 3) related to breast cancer risk or cognitive phenotype. One hundred and sixty-three SNPs, representing the 25 candidate genes, were identified in total. With the assistance of the University of Pittsburgh School of Nursing Molecular Genetics Laboratory manager, the student designed genotyping assays for selected SNPs using the Sequenom® iPLEX® MassARRAY platform (University of Pittsburgh Genomics and Proteomics Core Laboratories). One hundred

and forty-nine SNPs were amenable to inclusion in the iPLEX® MassARRAY. The student prepared 96 well plates of diluted DNA for processing.

Fourteen SNPs were not amenable to multiplexing. Of these 14 SNPs, four were deemed essential (due to functional consequence, location within the gene, or lack of alternative SNPs within a given gene) for inclusion. TaqMan® allelic discrimination assays were ordered for these four SNPs. All were successfully genotyped by the student using TaqMan® allelic discrimination in the School of Nursing Molecular Genetics Laboratory.

The multiplex results were received from the Core Laboratories and organized by the student. Thirteen SNPs included on the iPLEX® MassARRAY had call rates <90%. Of these 13 SNPs, five were deemed essential (due to functional consequence, location within the gene, or lack of alternative SNPs within a given gene). TaqMan® allelic discrimination assays were ordered for these five SNPs. Out of the five SNPs, two were successfully genotyped using TaqMan® allelic discrimination in the School of Nursing Molecular Genetics Laboratory. The remaining SNPs were not successfully genotyped using TaqMan® allelic discrimination. These SNPs were also not amenable to genotyping using a restriction fragment length polymorphism (RFLP) approach. Consequently, alternative SNPs were selected. Two of the substitute SNPs were successfully genotyped using TaqMan® allelic discrimination assays, while the third was successfully genotyped using RFLP.

Out of all of the 145 SNPs that were successfully genotyped, 14 were excluded due to minor allele frequencies in our sample of less than 0.05. In total, 131 SNPs are available to be included in the genetic analysis. Please refer to Figure 3 for a depiction of the genotyping workflow for this dissertation project.

2.2 PROPOSAL CHANGES

Several changes were made to the dissertation proposal and approved by the dissertation committee members at the student's comprehensive examination and overview. These changes, along with the rationale for these changes, are provided below.

2.2.1 Focus on pretreatment cognitive function assessment time point

The dissertation study originally proposed to investigate CTFs and DNA variation for associations with pretreatment cognitive function performance as well as cognitive decline with treatment at the first two assessment time points after initiation of systemic adjuvant therapy. Due to the fact that pretreatment findings were the impetus for this work and because of the hypothesis driving this study (i.e., disease-related factors inherent in breast cancer and/or host characteristics that predispose an individual to cancer as well as cognitive dysfunction may be a major determinant of cognitive changes in women with breast cancer), it was decided that all efforts related to this dissertation study should focus on the pretreatment assessment time point. Evaluation of CTFs and DNA variation for associations with treatment-related cognitive decline will be completed in a future study.

2.2.2 Omission of chemotherapy only treatment group

Originally, the dissertation study proposed to include a cohort of women with breast cancer prescribed/receiving chemotherapy alone. However, due to the limited sample size available for this cohort, in combination with the potential differences in the biology of breast cancers not

treated with anti-estrogen therapies, the dissertation committee members favored omission of the chemotherapy alone cohort from the data analysis.

2.2.3 Statistical analysis

Several changes have been made to the proposed statistical analysis as well. Use of both robust multiple linear regression models and standard multiple linear regression models eliminating potentially influential points to evaluate robustness of findings as part of a sensitivity analysis was originally proposed; however, due to the very large number of models, potential for error in eliminating individual points, improved model properties, and consistency in reporting results, the student, committee statistician, and committee chair made the decision to use a robust multiple linear regression model approach.

Because of the modification to focus on the pretreatment assessment time point, no multiple linear regression modeling at T1 or T2 will be performed for Aims 1, 2, or 3. Likewise, no linear mixed-effects regression modeling will be performed to evaluate the influence of CTF variation, DNA variation, and accompanying interactions over time for Aims 1, 2, or 3. Additionally for Aim 1, treatment cohort will no longer be included in models, and interactions between treatment cohort and CTF will not be considered due to the elimination of the longitudinal analyses.

Finally, since the original proposal, much work has been completed by the student and her committee on the development of a method to calculate and evaluate genetic risk/protection scores (GRSs) related to Aim 2. Weighted GRS will be calculated from regression coefficients from individual SNP models to assign greater risk/protection to SNPs with stronger associations. Overall GRS will be calculated across all candidate genes. Please refer to Appendix J to review

how our method was used to evaluate the collective effect of multiple oxidative stress and DNA repair variants on pretreatment cognitive function performance in postmenopausal women with breast cancer. In addition, interactions between GRSs and individual CTFs will be generated, instead of interactions between individual SNPs and CTFs, as originally proposed, to evaluate interaction effects as part of Aim 3.

2.3 STUDY STRENGTHS AND LIMITATIONS

This hypothesis-driven dissertation project is the first study to formally examine the impact of variation in CTFs of breast cancer and host DNA variation in genes used to clinically evaluate the biology of breast cancer on cognitive performance in postmenopausal women with early-stage breast cancer. Additional strengths of this dissertation project include the following: selection of biologically plausible candidate genes, representation of the variability in candidate genes through inclusion of both functional and tagging polymorphisms, well-characterized cognitive function phenotypes assessed with a battery of reliable and valid neuropsychological tests, adjustment for potential covariates and confounders of cognitive function performance, consideration of breast cancer heterogeneity through use of prescribed treatment group and inclusion of a matched healthy control group in the candidate gene analysis, and evaluation of the collective effect of possession of multiple risk or protection polymorphisms using weighted GRSs.

Study limitations should also be acknowledged. Small samples sizes limited the interpretability of results from the CTF-by-GRS interaction analysis and the full clinicopathologic and genetic predictive model analysis as well as our ability to evaluate gene-

dosage effects. The study sample was comprised of primarily white postmenopausal women with hormone receptor positive early-stage breast cancer; the generalizability of our findings to more diverse populations and breast cancers is unknown. For CTF data, we were limited to information available in surgical pathology reports obtained from the medical record. In addition, we did not account for the potentially confounding effects of surgery and/or anesthesia exposure on pretreatment cognitive performance in our analysis.

2.4 FUTURE STUDIES AND IMPLICATIONS FOR NURSING

Findings from this dissertation study need to be confirmed in larger, more diverse populations and cancers. Ideally, future studies will include a cognitive function assessment prior to primary surgery as a means to more fully capture the influence of breast cancer heterogeneity, and subsequent tumor removal, on cognitive performance. Future analyses should also investigative the effect of variation in CTFs of breast cancer and host DNA in genes used to clinically evaluate the biology of breast cancer, on cognitive function throughout and following completion of adjuvant therapy. Because candidate genes evaluated as part of this dissertation project were identified from prognostic multigene expression profiles for breast cancer and many of our significant findings were related to polymorphisms with functional consequences and/or located in regulatory regions of genes, tumor gene expression levels should be prioritized as an additional focus of investigation.

The findings from this dissertation project inform current knowledge related to biological underpinnings for pretreatment cognitive dysfunction in women with breast cancer and provide the foundation for a line of research to identify clinically relevant biomarkers and novel therapies

for breast cancer survivors experiencing cognitive dysfunction. We envision a future where women with breast cancer will not only receive a refined breast cancer diagnosis based on clinicopathologic and genetic characteristics but also tailored symptom prediction and proactive symptom management. Nurses will be at the forefront of patient education related to symptom prediction, administration and coordination of biomarker testing, and creation of holistic, patient-centered care plans that feature interventions intended to help mitigate negative cancer- and cancer treatment-related cognitive symptoms.

3.0 DATA-BASED MANUSCRIPT: THE IMPACT OF VARIATION IN CLINICOPATHOLOGIC TUMOR FEATURES AND BREAST CANCER-RELATED GENETIC POLYMORPHISMS ON PRETREATMENT COGNITIVE FUNCTION IN WOMEN WITH BREAST CANCER: AN EXPLORATORY ANALYSIS

3.1 ABSTRACT

Breast cancer is a heterogeneous disease characterized by molecular and pathologic diversity. Based on previous work, we hypothesized that the cellular heterogeneity in breast cancers may account for variability in the presence and/or severity of cognitive dysfunction among women diagnosed with breast cancer, especially prior to initiation of systemic adjuvant therapy. The purpose of this study was to investigate relationships between clinicopathologic tumor features (CTFs) of breast cancer or host DNA variation in genes used to clinically evaluate the biology of breast cancer and cognitive performance in postmenopausal women with early-stage breast cancer. Interactions between CTFs and DNA variation were also explored. Pretreatment cognitive function assessment occurred after surgery but before initiation of systemic adjuvant therapy. CTF data were obtained from surgical pathology reports for women with breast cancer scheduled to receive adjuvant anastrozole therapy ± chemotherapy (n=329). Genotypes for 131 functional/tagging single nucleotide polymorphisms (SNPs) related to 25 biologically plausible breast cancer-related candidate genes (*AURKA, BAG1, BCL2, BIRC5, CCNB1, CD68, CENPA, CMC2, CTSL2, DIAPH3, ERBB2, ESR1, GRB7, GSTM1, MELK, MKI67, MMP11, MYBL2, NDC80, ORC6, PGR, RACGAP1, RFC4, RRM2,* and *SCUBE2*) were determined for three groups of women: women with breast cancer, prescribed chemotherapy followed by anastrozole (n=55); women with breast cancer prescribed anastrozole only (n=83); and postmenopausal age- and education-matched controls without cancer (n=82). Standard and robust multiple linear regression models were used to determine if CTFs and SNPs accounted for variability in cognitive performance scores. Weighted multi-gene, multi-SNP genetic risk/protection scores (GRSs) were calculated based on significant ($p<0.05$) individual SNP results as a means to evaluate the collective effect of multiple SNPs on cognitive performance. Interactions between

CTFs and GRSs were also assessed using linear regression models. Significant ($p<0.05$) relationships were reported between cognitive performance on one or more cognitive function composites and the following CTFs: cancer stage; tumor size; tumor focality; tumor laterality; tumor location octant; invasive type; Nottingham Score; Nottingham Grade; estrogen receptor (ER) H-score; progesterone receptor (PR) status; HER2 status; HER2 immunohistochemistry classification score; Ki67 classification; Ki67 index; Oncotype DX® Breast Cancer Assay Recurrence Score®; Oncotype DX® Breast Cancer Assay quantitative single gene ER, PR, and HER2 scores; and Magee Equation recurrence score. Significant ($p<0.05$) SNP main effect and/or SNP-by-prescribed treatment group interactions were observed between at least one cognitive function composite and one or more SNPs of all candidate genes evaluated with the exception of *CMC2*, *MMP11*, and *RACGAP1*. Each computed GRS was found to be significantly ($p<0.001$) and positively (i.e., as overall genetic protection increases, cognitive performance score increases indicating better performance) associated with its corresponding cognitive function composite score. Only one significant interaction, between tumor location octant and visual working memory GRS, was noted. Overall, results from this exploratory study suggest that clinicopathologic and breast cancer-related genetic variation influence pretreatment cognitive performance in women with breast cancer and merit additional investigation.

Keywords: breast neoplasms; cognition; genetics; pathology; pretreatment

3.2 BACKGROUND

Breast cancer is not a single disease, but rather a heterogeneous collection of diseases characterized by high degrees of molecular and pathologic diversity both among breast cancers

diagnosed in different individuals (i.e., inter-tumor) and within the same breast tumor (i.e., intra-tumor) (Martelotto, Ng, Piscuoglio, Weigelt, & Reis-Filho, 2014; Polyak, 2011; Rivenbark, Connor, & Coleman, 2013). Despite the well-established inter-tumor and intra-tumor heterogeneity in breast cancers, few studies have examined how variation in the biology of cancer itself may impact the presence and/or severity of cognitive symptoms experience by many women diagnosed with breast cancer (Ahles & Saykin, 2007; Ahles, Root, & Ryan, 2012).

Previous investigations have focused on the direct neurotoxic effects of adjuvant chemotherapy and the anti-estrogen effects of adjuvant aromatase inhibitors and selective estrogen receptor modulators on cognitive performance in breast cancer survivors over time with treatment (Ono et al., 2015; Zwart, Terra, Linn, & Schagen, 2015). However, as succinctly summarized in a recent review by Wefel, Kesler, Noll, and Schagen, a number of studies have also reported evidence of cognitive dysfunction in women diagnosed with breast cancer prior to initiation of adjuvant chemotherapy and/or anti-estrogen therapy, unrelated to distress, fatigue, comorbidities, or surgery-related factors (2015).

While multiple pathophysiological mechanisms likely underlie cognitive changes in women with breast cancer, these pretreatment findings support the hypothesis that disease-related factors inherit to breast cancer, including clinicopathologic tumor features (CTFs) of breast cancer and host DNA variation in genes used to clinically evaluate the biology of breast cancer, may contribute to cognitive dysfunction symptomatology and warrant more in depth investigation (Ahles & Saykin, 2007; Ahles et al., 2012; Wefel et al., 2015). We further hypothesize that the heterogeneity in disease-related factors of breast cancer at the cellular level may account for a significant proportion of observed variability in cognitive dysfunction among

women diagnosed with breast cancer, especially prior to initiation of systemic adjuvant therapy (Koleck & Conley, 2016).

Although limited, evidence exists to support the investigation of CTFs and a genetic component of cognitive dysfunction in women with breast cancer. In relation to CTFs, Ahles and colleagues found that at pretreatment, women diagnosed with invasive early-stage breast cancer (Stages 1-3) had significantly poorer reaction time than healthy controls; in contrast, participants diagnosed with noninvasive cancer (Stage 0) did not perform more poorly than healthy controls (2011). Mandelblatt and colleagues reported that older women (≥ 60 years of age) diagnosed with Stage 2-3 breast cancer had lower pretreatment executive function scores compared to participants with Stage 0-1 breast cancer (2014). In support of the genetic component of cognitive dysfunction, four published studies have reported associations between *APOE* genotype and cognitive performance in women with breast cancer (Ahles et al., 2011; Ahles et al., 2003; Koleck et al., 2014; Lengacher et al., 2015). Associations with polymorphisms in *ANKK1*, *BDNF*, *COMT*, *MTHFR*, and *SLC64A* have also been reported (Lengacher et al., 2015; Ng et al., 2015; Small et al., 2011).

However, to our knowledge, no study to date has formally investigated the mechanistic merit of the previously stated hypotheses. Therefore, the purpose of the current exploratory study is to investigate how heterogeneity in breast cancers at the molecular and pathologic levels is related to reported variability in pretreatment neuropsychological performance in postmenopausal women diagnosed with early-stage breast cancer. Specifically, the three distinct, but interrelated, aims of this study are as follows: (1) to investigate the relationship between CTFs of breast cancer, evaluated as part of standard surgical pathology reports, and cognitive function in postmenopausal women with early-stage breast cancer; (2) to explore DNA

variation in genes used to clinically evaluate the biology of breast cancer, as part of prognostic multigene expression profiles for breast cancer, for association with susceptibility to or protection from cognitive decline in postmenopausal women with early-stage breast cancer; and (3) to explore interactions between CTFs of breast cancer and DNA variation in genes used to clinically evaluate the biology of breast cancer on cognitive function.

3.3 METHODS

3.3.1 Study sample

Participants in this exploratory, ancillary study were originally recruited as part of a larger, ongoing parent study investigating the effects of adjuvant chemotherapy and the adjuvant anti-estrogen therapy, anastrozole, on changes in cognitive function in postmenopausal women with breast cancer (Bender et al., 2015). The current study analyzes three interrelated subgroups of participants: (1) the subgroup for the CTF analysis was comprised of 329 postmenopausal women newly diagnosed with Stage 1, 2, or 3a breast cancer, with no evidence of metastases, scheduled to receive either chemotherapy followed by anastrozole or anastrozole only; (2) the subgroup for the genetic analysis (N=220) included the cohort of women from subgroup #1 who provided a specimen for genetic evaluation (n=138) in addition to postmenopausal age-and education-matched healthy controls without breast cancer (n=82) who also provided a genetic specimen; and, lastly, (3) the subgroup for the interaction analysis was limited to women with breast cancer with both CTF and genetic data available.

All study participants were postmenopausal, 75 years of age or younger, able to speak and read English, and completed a minimum of eight years of education. Participants were excluded from the parent study, and consequently this ancillary study, if they had a prior history of neurologic disease or cancer or were hospitalized for psychiatric illness within the past two years. Informed consent was obtained from all study participants. Both the ancillary and parent studies were approved by the University of Pittsburgh Institutional Review Board.

3.3.2 Evaluation of cognitive function

Cognitive function was evaluated using a battery of 13 reliable, validated neuropsychological measures intended to assess multiple cognitive domains. Rationale for selection of individual tests has been described previously (Bender et al., 2015). Women in the breast cancer cohorts completed the cognitive function test battery after primary surgery but before initiation of systemic adjuvant therapy. Healthy women completed the same cognitive function test battery. Tests were administered to participants and scored by trained research nurses. Due to the number of cognitive measures, an exploratory factor analysis with principal component extraction was previously employed to reduce the dimensionality of the cognitive function data (Bender et al., 2015). This data reduction technique identified eight cognitive function composites: attention [Cambridge Neuropsychological Test Automated Battery (CANTAB) Rapid Visual Information Processing Test (Robbins et al., 1994)], concentration [Digit Vigilance Test (Lafayette Clinical Insturments Company, 1989)], mental flexibility [Delis Kaplan Executive Function System Color-Word Interference Test (Delis et al., 2001)], executive function [CANTAB Stockings of Cambridge Test (Robbins et al., 1994) and CANTAB Spatial Working Memory Test (Robbins et al., 1994)], psychomotor speed [Grooved Pegboard Test

(Klove, 1963) and Digit Symbol Substitution Test (Wechsler, 1998)], verbal memory [Rey Auditory Verbal Learning Test (Rey, 1941), Delis Kaplan Executive Function Verbal Fluency Test (Delis et al., 2001), and Rivermead Story Test (Cockburn & Smith, 1993)], visual memory [CANTAB Paired Associates Learning Test (Robbins et al., 1994) and Rey Complex Figure Test (Osterrieth, 1944)], and visual working memory [CANTAB Stockings of Cambridge Test (Robbins et al., 1994) and Rey Complex Figure Test (Osterrieth, 1944)]. The eight cognitive function composites are utilized as the outcome variables for this study. Participant cognitive function composite z-scores were calculated such that more positive scores are associated with better cognitive performance and more negative scores are associated with poorer cognitive performance.

3.3.3 Assessment of potential covariates and confounders

Potential covariates and confounders of cognitive function were assessed in all cohorts of study participants at the same time as cognitive function assessment and included: age (in years); estimated verbal intelligence – National Adult Reading Test-Revised (Nelson, 1982), depressive symptoms – Beck Depression Inventory-II (Beck et al., 1996); anxiety – Profile of Mood States Tension-Anxiety subscale (McNair et al., 1992); fatigue – Profile of Mood States Fatigue-Inertia subscale (McNair et al., 1992); and current pain – Brief Pain Inventory (Cleeland, 1989). Only participants with complete covariate/confounder information were included in analyses.

3.3.4 Evaluation of CTFs

CTF data were obtained from surgical pathology reports of study participants and included the following: TNM Classification of Malignant Tumors overall stage (Stage 1, 2a, 2b, or 3a) (Edge et al., 2010); TNM Classification of Malignant Tumors tumor stage (Stage T1a, T1b, T1c, T2, T3) (Edge et al., 2010); lymph node status (positive or negative); number of positive lymph nodes; tumor site laterality (left or right breast); tumor location within breast [clock position and/or quadrant location (upper outer, upper inner, lower outer, or lower inner)]; tumor focality (single or multiple); primary tumor size (measured to the nearest millimeter); aggregate tumor size if multifocal (measured to the nearest millimeter); histologic type (invasive ductal, invasive lobular, or both) (Tavassoli & Devilee, 2003); combined histologic Nottingham Score [score 3-9; sum of three subscores: glandular/tubular differentiation score (1-3), nuclear pleomorphism score (1-3), and mitotic activity/count score (1-3)] (Elston & Ellis, 1991); combined histologic Nottingham Grade [Grade 1 (low), Grade 2 (intermediate), or Grade 3 (high)] (Elston & Ellis, 1991); lymphovascular invasion (presence or absence); estrogen receptor (ER) status (positive or negative); ER H-score (extent of nuclear immunoreactivity) quantitation score (score 0-300); ER Oncotype DX® Breast Cancer Assay quantitative single gene score (score 0-12.5+; Negative <6.5, Positive ≥6.5) (Genomic Health Inc., 2016); progesterone receptor (PR) status (positive or negative); PR H-score (extent of nuclear immunoreactivity) quantitation (score 0-300); PR Oncotype DX® Breast Cancer Assay quantitative single gene score (score 0-10+; Negative <5.5, Positive ≥5.5) (Genomic Health Inc., 2016); HER2 immunohistochemistry (IHC) classification score [0, 1+ (Negative), 2+ (Equivocal), or 3+ (Positive)] (Wolff et al., 2007); HER2/neu status (positive or negative based on IHC test and/or FISH amplification); HER2 Oncotype DX® Breast

Cancer Assay quantitative single gene score (score 0-13+; Negative <10.7, Equivocal 10.7-11.5, Positive ≥11.5) (Genomic Health Inc., 2016); Ki67 index (0-100%; percentage of total number of tumor cells with nuclear staining); Ki67 proliferative rate classification [Low (≤10%), Moderate (11-25%), High (26-50%), or Very High (>50%)]; and Oncotype DX® Breast Cancer Assay Recurrence Score® (score 0-100 from multigene expression algorithm) (Paik et al., 2004, 2006). In instances where a participant had more than one primary breast tumor in the same breast or bilateral breast cancer, characteristics of the tumor contributing to the highest overall breast cancer stage were used in analyses.

Please note that clock position and quadrant information collected to describe the location of the tumor within the breast was combined and condensed into eight tumor locations: upper outer, upper inner, lower outer, lower inner, upper middle (12 o'clock), lower middle (6 o'clock), outer middle (left breast-3 o'clock; right breast-9 o'clock), and inner middle (left breast-9 o'clock; right breast-3 o'clock). Due to limited numbers, retroareolar tumors were classified by clock position, if provided, or omitted from the analysis. In the case of multifocal tumors with foci identified in multiple quadrants, the average clock position was used if continuous between foci or the case was omitted from the analysis. These eight locations are referred to as tumor location *octants* throughout the text.

As a supplement to Oncotype DX® Breast Cancer Assay Recurrence Scores®, Magee Equation recurrence scores were calculated using the three equations described in Klein et al. (2013). The three equations, which produce very similar results, use different combinations of Nottingham Score, ER H-score, PR H-score, HER2 status (negative, equivocal, or positive), tumor size, and/or Ki67 index to estimate Oncotype DX® Breast Cancer Assay Recurrence Scores® and corresponding recurrence risk category assignment (i.e., low, intermediate, or high).

Thus, up to three scores were calculated for each participant based on available information. Scores from the three equations were reduced into a single variable giving preference to generated scores in the following sequence: equation 1 > equation 2 > equation 3. Scores from equation 1 were prioritized as this equation was found to most accurately replicate extreme values (i.e., assignment into the low and high recurrence risk categories). If a score from equation 1 was not available for a participant, the score from equation 2 was selected due to its concordance with Oncotype DX® Breast Cancer Assay risk category overall and comparable performance to equation 3 when the intermediate risk category was omitted.

3.3.5 Candidate gene selection and genotype data collection

In total, 25 biologically plausible candidate genes (*AURKA, BAG1, BCL2, BIRC5, CCNB1, CD68, CENPA, CMC2, CTSL2, DIAPH3, ERBB2, ESR1, GRB7, GSTM1, MELK, MKI67, MMP11, MYBL2, NDC80, ORC6, PGR, RACGAP1, RFC4, RRM2,* and *SCUBE2*), that are theorized to represent the biology of breast cancer at the cellular level, were selected for investigation. The rationale behind the identification and prioritization of candidate genes from multigene expression profiles for breast cancer has been described previously (Koleck & Conley, 2016). Briefly, because multigene expression profiles for breast cancer enhance knowledge received from traditional CTFs via algorithm-driven estimation of adjuvant therapy benefit and risk of distant metastasis or recurrence, genes from these profiles play an important role in breast cancer aggressiveness and progression, complement the collected CTF data, and are ideal for exploring the overall study hypotheses.

Single nucleotide polymorphisms (SNPs) representing each candidate gene were identified. Functional polymorphisms within, directly upstream of, and/or found to influence

expression levels of candidate genes, were identified from the literature. When a functional polymorphism was not identified and/or did not fully represent the gene of interest, tagging SNPs were selected using the Phase III HapMap database. Because the profiles from which candidate genes were selected rely upon gene expression data, evaluation of DNA variability was extended ±2,500bps beyond the gene to capture the UTR-5' and UTR-3' regulatory regions. Criteria for selection of tagging SNPs were as follows: R^2 of ≥ 0.8, minor allele frequency (MAF) of ≥ 0.20, and selected for Caucasian ancestry, which represents the majority of study participants. One-hundred and sixty-three functional and tagging SNPs were identified to represent the variability in the 25 candidate genes.

Genetic samples were collected from participants between June 2008 and May 2014. Three milliliters of whole blood or two milliliters of saliva were obtained for genotyping. DNA was extracted from peripheral blood leukocytes using a simple salting out procedure (Miller, Dykes, & Polesky, 1988) or from saliva following the protocol and reagents supplied with the Oragene® DNA collection kits (DNA Genotek Inc., 2012). The Sequenom® iPLEX® MassARRAY platform (Sequenom, San Diego, CA) was used as the primary genotyping method for this study due to the number of SNPs and participants evaluated. SNPs were also genotyped using TaqMan® allele discrimination with the ABI Prism 7000 Sequence Detection System (SDS) and SDS software v1.2.3 (Thermo Fisher Scientific Inc., Waltham, MA) as well as a restriction fragment length polymorphism approach.

Negative controls were included with all analyses. Genotypes were double called by individuals blinded to participant phenotypes and discrepancies were addressed by reviewing raw data or re-genotyping. Participant genotypes were classified for data analysis based on the

presence or absence of the minor allele (homozygous wild type compared to the combination of heterozygotes and homozygous variant genotypes).

3.3.6 Data cleaning and quality assurance

Collected CTF data were independently entered into a computer database by two individuals blinded to outcome and genetic data and compared for discrepancies. Discrepancies were adjudicated by a third individual by independent review of raw data. SNPs with call rates less than 90% or MAFs of less than 0.05 were omitted. For SNPs not meeting the 90% call rate threshold but deemed essential for inclusion in the study (due to functional consequence, location within a candidate gene, or lack of alternative SNPs available within a given gene) secondary genotyping approaches were attempted. Alternatives were selected for essential SNPs in instances of multiple failed genotyping attempts and/or lack of availability of alternative genotyping methods.

Furthermore, detailed data screening procedures were performed to ensure data accuracy. Data from individual CTFs were cross-checked with other, directly corresponding CTFs (e.g., tumor stage and tumor size). Inconsistencies were addressed by reviewing raw data from CTF collection forms and/or the original pathology reports. Each SNP was tested for Hardy-Weinberg Equilibrium (HWE), using chi-square goodness-of-fit or Fisher's exact tests, to identify potential genotyping errors.

3.3.7 Statistical analysis

Analyses were performed using Stata® Data Analysis and Statistical Software SE Version 14.1 (StataCorp, College Station, TX) and IBM® SPSS® Statistics Version 23 (IBM Corp., Armonk, NY). Standard descriptive statistics were computed for all predictors, outcomes, and covariates/ confounders. Separate overall participant demographic, covariate/confounder, and cognitive function composite z-score summary statistics were calculated for both the CTF and genetic data analyses. Overall CTF data were summarized using means, standard deviations, minimum and maximum values, frequencies, and percentages. Additionally, due to findings from the individual CTF regression analyses related to tumor location octant, tumor characteristics were compared by tumor location octant using one-way ANOVAs to compare means of continuous variables and Fisher's exact tests, computed using 2-sided Monte Carlo sampling based on 10,000 sampled tables, to examine associations between categorical variables. For the genetic analysis, differences in demographics, covariates/confounders, cognitive function composite z-scores, and CTFs by study cohort were evaluated using one-way ANOVAs and Pearson's chi-square tests of independence (when expected cell counts ≥ 5), Fisher's exact tests (for 2×2 contingency tables with expected cell counts ≤ 5), or Fisher's exact tests computed using 2-sided Monte Carlo sampling based on 10,000 sampled tables (for contingency table larger than 2×2 with expected cell counts ≤ 5). The comparability of covariate/confounder data and cognitive function performance z-scores between participants included in the CTF and genetic analyses and remaining participants from the parent study were also assessed.

For the CTF analysis, standard multiple linear regression models and complementary robust multiple linear regression models using Huber weighting and biweighting iterations were fit to estimate associations between individual CTFs and each cognitive function composite

score. In order to evaluate potential nonlinear (e.g., quadratic) relationships between continuous CTFs and cognitive function composites, squared versions of all continuous CTFs were computed. Both standard and robust linear regression models with the original and squared version of a particular CTF were produced. Furthermore, all two-way interactions between marginally significant ($p<0.20$) CTF predictors, identified via the robust regression models with linear and squared terms, were evaluated using standard and robust multiple linear regression modeling. In cases of significant squared predictors, interactions were generated with both the original and squared variable and all terms were included in the model. All models controlled for age, estimated intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Underlying assumptions were assessed for each regression model. Specifically, histograms of jackknifed residuals were used to evaluate normality, added variable plots were generated to assess linearity, Breusch-Pagan/Cook-Weisberg tests were conducted to evaluate homoscedasticity, and variance inflation factors were used to assess for multicollinearity. In order to identify potentially influential points, Cook's distance was generated and evaluated as part of jackknifed residual by predicted value scatterplots. Due to concerns related to influential points and heteroscedasticity, robust regression model estimated regression coefficients and corresponding significance levels are reported.

Standard and robust multiple linear regression modeling was also used to examine relationships between each cognitive function composite and the presence (i.e., homozygous variant genotype plus heterozygous genotype) or absence (i.e., homozygous wildtype genotype) of one or more minor alleles for each SNP. In order to account for the heterogeneity of breast cancer tumors in the genetic analysis, women diagnosed with breast cancer were further classified using prescribed future treatment regimen as a surrogate for disease characteristics.

Thus, the genetic analysis included two groups of women diagnosed with breast cancer, those prescribed chemotherapy followed by anastrozole (n=55) and those prescribed anastrozole only (n=83), as well as the matched healthy control group (n=82). Both main SNP effects only and SNP-by-prescribed treatment group interaction effect regression models were fitted. In all models, healthy controls served as the reference group for the two prescribed treatment groups (i.e., prescribed chemotherapy plus anastrozole or anastrozole alone). Likewise, the wildtype genotype served as the reference group for possession of one or more minor alleles. All regression models were adjusted for age, estimated intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain, and prescribed treatment group. Underlying assumptions were assessed for each regression model using the same techniques described for the CTF analysis. Again, to lessen the impact of potentially influential points and adjust for heteroscedasticity, robust regression model estimated regression coefficients and significance levels are reported.

Genetic risk/protection scores (GRSs) for each cognitive function composite were then calculated to explore the influence of possession of multiple significant ($p<0.05$) genotypes on cognitive function composite scores. SNP minor alleles that were significantly ($p<0.05$) negatively or positively associated with a cognitive function composite by either SNP main effects and/or SNP-by-prescribed treatment interaction effects were used in GRS calculations. A weighted calculation method, in which unstandardized robust regression coefficients from the individual genetic models were multiplied by 0 (absence) or 1 (presence) based on a participant's genotype and prescribed treatment group membership and then summed, was used to assign greater risk/protection to minor alleles with stronger associations. A lower GRS conveys greater genetic risk for poorer cognitive function and a higher GRS conveys greater genetic protection.

GRSs were added as the final predictor to standard and robust multiple linear regression models adjusted for age, estimated verbal intelligence, levels of depressive symptoms, anxiety, fatigue, pain, and prescribed treatment group. Only participants with all genetic data necessary for calculation of a particular GRS were included in the GRS analysis.

A genetic sub-analysis, featuring only women with breast cancer, was also completed for *ERBB2* and *MKI67*. Because *ERBB2* and *MKI67* have rare or limited expression in normal breast tissue (Pavelic et al., 1992; Stark et al., 2000; Urruticoechea, Smith, & Dowsett, 2005), we hypothesized that polymorphism in these genes may only be impactful if the genes are being expressed. Thus, standard and robust multiple linear regression analyses for *ERBB2* and *MKI67* SNPs by HER2 IHC classification score or Ki67 index, respectively, instead of prescribed treatment group, was conducted. Models with HER2 IHC classification score, an *ERRB2* SNP, and SNP-by-HER2 IHC classification interaction as predictors were fitted for each *ERBB2* polymorphism. Likewise, models with Ki67 index, an *MKI67* SNP, and SNP-by-Ki67 index interaction as predictors were fitted for both *MKI67* polymorphisms. All models controlled for age, estimated intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain. Robust regression model regression coefficients and p-values are reported.

For each cognitive function composite, interactions between significant ($p<0.05$) individual CTFs, identified from models with linear and quadratic terms, were tested for interactions with calculated GRSs using standard and robust multiple linear regression modeling. Interaction terms were created as the products of individual CTFs and GRSs. Only participants with breast cancer and genetic data were included in the interaction analysis. In instances of significant squared CTF predictors, interactions were generated between the GRS and both the

original and squared variable; all terms were included in the model. All models were adjusted for age, estimated intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Full standard and robust multiple linear regression models were created to explore the potential to better predict cognitive function performance using a combination of both CTF and genetic data. First, a final CTF model was developed for each cognitive function composite. Statistically significant ($p<0.05$) terms from individual main effect and interaction effect models were evaluated by placing all predictors into a single robust regression model. Predictors were omitted from the model if multicollinearity and/or limited variability were present. All predictors meeting the screening criterion ($p<0.20$) were included in the final CTF models. GRSs were then added as a predictor to the final CTF models. All models were adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain. R^2s obtained from standard regression models are reported.

3.4 RESULTS – CLINICOPATHOLOGIC TUMOR FEATURES

3.4.1 Participant and breast cancer tumor characteristics

Of the 354 women diagnosed with early-stage breast cancer from which CTF data were collected, a total of 329 had complete confounder/covariate information and cognitive function scores available for one or more cognitive function composites. A summary of overall demographic, covariate/confounder, and cognitive function data for participants included in this analysis is located in Table 7. In general, participants were an average of 61.05 years of age, well-educated (with a mean of 14.80 years of education), married or currently living with a

partner (67.8%), and Caucasian (96.4%). A comparison of characteristics of participants included (n=329) to those not included (because CTF or pretreatment cognitive function data were not available or covariate/confounder information was incomplete) (n=40) revealed that participants not included in the analysis had poorer (p=0.003) mean attention performance z-scores (-0.66±1.162) than participants included in the analysis (-0.16±0.939).

The majority of breast cancer tumors were ductal (86.9%), single focus (84.2%), overall Stage 1 (65%), tumor stage T1c (40.4%), lymph node negative (77.5%), ER positive (98.8%), PR positive (87.8%), and HER2 negative (91.2%). The mean Nottingham Score (6.04±1.306) for all tumors included in the analysis corresponds to an intermediate Nottingham Grade, and the mean Ki67 index (23.10±21.522) reflects a moderate Ki67 classification. Oncotype DX® Breast Cancer Assay Recurrence Scores® ranged from 0 to 63 with a mean score of 18.26±9.76. Similarly, Magee Equation recurrence scores ranged from 1.92-48.87 with a mean score of 20.51±7.77. A comprehensive summary of all CTF data included in this analysis is reported in Table 8.

3.4.2 Individual CTFs and cognitive function

Table 9 reports the regression coefficients and p-values from all robust regression models evaluating the relation between individual CTFs and cognitive function composites. Table 10 complements the results from Table 9 by considering nonlinear quadratic relations and presents regression coefficients and p-values from robust regression models with squared continuous CTF terms.

The most significant findings were related to memory and HER2 status and HER2 IHC classification score. Possession of a HER2 positive tumor contributed to poorer verbal (b=-

0.287, p=0.018), visual (b=-0.270, p=0.001), and visual working (b=-0.490, p<0.001) memory performance compared to possession of a HER2 negative tumor. Likewise, as HER2 IHC classification scores increased, verbal (b=-0.072, p=0.093), visual (b=-0.081, p=0.003), and visual working (b=-0.170, p<0.001) memory performance scores decreased (Figure 4). Moreover, when considering a quadratic continuous predictor model, a significant relationship between visual memory performance and Oncotype DX® Breast Cancer Assay HER2 quantitative single gene score (b=1.983, p<0.001) and the corresponding squared term (b=-0.104, p<0.001) was observed.

In addition to associations with HER2 status and HER2 IHC score, a significant association was noted between tumor focality and verbal memory (b=-0.278, p=0.003), such that possession of a multifocal tumor contributed to poorer performance compared to a single focus tumor. While not statistically significant, this trend can be seen across all cognitive function composites. Possession of a progesterone receptor (PR) positive tumor, compared to a PR negative tumor, also contributed to poorer verbal memory performance (b=-0.256, p=0.015).

Intriguing tumor location effects were also noted. To begin, having a tumor located in the left breast, compared to the right breast, contributed positively to verbal memory (b=0.156, p=0.025) and visual working memory (b=0.163, p=0.026) performance scores. Overall tumor location octant was found to be significantly (p=0.027) related to mental flexibility. Specifically, having a tumor in the lower inner octant contributed to poorer mental flexibility performance compared to having a tumor in the upper outer octant (b=-0.511, p=0.001). While not statistically significant at the p<0.05 level, tumors located in the lower outer (b=-0.207, p=0.145), upper inner (b=-0.171, p=0.147), upper middle (b=-0.180, p=0.121), or lower middle (b=-0.258, p=0.100) octants also appear to contribute to poorer mental flexibility performance

compared to the upper outer octant. Although tumor location octant did not significantly contribute to the model as a whole, the contribution of the individual upper inner octant to attention performance (b=-0.315, p=0.040) and the lower middle octant to visual working memory performance (b=-0.432, p=0.007) significantly differed compared to that of the upper outer octant.

Additionally, as Oncotype DX® Breast Cancer Assay Recurrence Score® increased, mental flexibility performance score decreased (b=-0.010, p=0.032; Figure 5). In the quadratic continuous predictor models, significant relations were found between Oncotype DX® Breast Cancer Assay Recurrence Score® and verbal memory performance (b=0.032, p=0.021; b=-0.001, p=0.029) as well as Magee Equation recurrence score and visual memory performance (b=0.026, p=0.055; b=-0.001, p=0.034) and visual working memory performance (b=0.051, p=0.028; b=-0.001, p=0.021).

Finally, Ki67 classification was found to be significantly associated with concentration performance (p=0.042). In particular, a moderate Ki67 classification contributed positively to cognitive function performance compared to a low Ki67 classification (b=0.381, p=0.009).

3.4.3 CTF differences by tumor location

In order to facilitate interpretation of the tumor location effects, an evaluation of how CTFs differed by tumor location octant was completed. Features that were found to be significantly ($p<0.20$) different by tumor location octant are presented in Table 11. Tumors differed by tumor size, aggregate tumor size, tumor focality, invasive type, Nottingham Grade, HER2 status, HER2 IHC classification score, and lymphovascular invasion presence. Of particular interest, tumors in the lower middle octant have a higher percentage of HER2 positive tumors (30.0%) and higher

mean HER2 IHC classification score (1.94±0.929) than the other octants. These findings connect three of our significant results: (1) lower middle octant visual working memory performance significantly differed compared to the upper outer octant; (2) possession of a HER2 positive tumor contributed to poorer visual working memory performance compared to possession of a HER2 negative tumor; and (3) as HER2 IHC classification scores increased, visual working memory performance scores decreased.

3.4.4 Two-way CTF interactions and cognitive function

All marginally significant ($p<0.20$) individual CTFs tested for two-way interactions are listed by cognitive function composite in Table 12. CTFs identified from both linear and quadratic robust regression models were included. All significant ($p<0.05$) two-way CTF interactions are displayed in Table 13. One or more significant two-way interactions were identified for each cognitive function composite.

3.5 RESULTS - CANDIDATE GENE ANALYSIS

3.5.1 Participant and tumor characteristics by study cohort

Genetic data was collected from 226 participants. Of these 226 participants, 220 had complete covariate/confounder information and cognitive function scores available for one or more cognitive function composites. A summary of overall demographic, covariate/confounder, and cognitive function data for participants included in this analysis can be found in Table 14.

Genetic analysis cohorts (i.e., prescribed chemotherapy plus anastrozole, prescribed anastrozole alone, and healthy controls) differed statistically, yet not clinically meaningfully, by age and estimated verbal intelligence (Table 15). The groups also differed by level of anxiety (p=0.003), with women with breast cancer prescribed chemotherapy plus anastrozole having higher mean anxiety levels (9.61±6.140) than women with breast cancer prescribed anastrozole only (6.97±4.654) and the healthy controls (6.55±5.619). Comparison of CTFs by prescribed treatment group confirmed differences in disease characteristics (Table 15). To summarize, women with breast cancer prescribed chemotherapy plus anastrozole had higher frequency of Stage 2a, 2b, and 3a breast cancers, larger mean tumor size, higher mean number of positive lymph nodes, higher mean Nottingham Score, greater frequency of lymphovascular invasion, lower ER H-score, lower Oncotype DX® Breast Cancer Assay quantitative single gene PR score, greater frequency of HER2 positive cancer, higher mean Ki67 index, higher mean Oncotype DX® Breast Cancer Assay Recurrence Score®, and higher mean Magee Equation recurrence score compared to women with breast cancer prescribed anastrozole alone.

No differences in covariates/confounders or pretreatment cognitive function z-scores were observed between healthy control women included in the genetic analysis and remaining parent study participants not included in the genetic analysis (n=82). Women with breast cancer prescribed anastrozole only included in the genetic analysis had slightly lower (p=0.044) mean estimated verbal intelligence (107.04±8.844) than those not included in the genetic analysis (n=155; 109.42±8.542). Women with breast cancer prescribed chemotherapy plus anastrozole included in the genetic analysis had higher (p=0.014, p=0.006, p=0.002, respectively) mean pretreatment verbal (0.02±0.662), visual (0.29±0.352), and visual working (0.30±0.514) memory

performance z-scores compared to those not included in the genetic analysis (n=78; -0.28±0.697, 0.03±0.615, -0.07±0.746, respectively).

3.5.2 Candidate gene SNP quality assurance

Of the 163 SNPs originally identified, 18 nonessential SNPs (that were not amenable to multiplexing or had call rates less than 90%) and 14 SNPs with study MAFs of less than 0.05 were excluded. Alternatives were selected for three essential SNPs. In total, 131 SNPs were included in the genetic analysis (Table 16). Genotyping call rates for these SNPs ranged from 90 to 100%. When considering all study participants, six SNPs were not in HWE: *CTSL2*rs4361859 (p=0.0078), *ESR1*rs2234693 (p=0.0344), *ORC6*rs33994299 (p=0.0051), *PGR*rs1042838 (p=0.0466), *PGR*rs1042839 (p=0.0103), and *PGR*rs474320 (p=0.0434). When considering the healthy control women alone, *PGR*rs1042838 (p=0.0160), *PGR*rs1042839 (p=0.0027), and *PGR*rs474320 (p=0.0329) still did not meet HWE. The deviation from HWE is most likely attributed to non-random sampling of study participants from the population.

3.5.3 Individual SNPs and cognitive function

Individual polymorphisms significantly ($p<0.05$) associated with a cognitive function composite by either SNP main effects or SNP-by-prescribed treatment group interaction effects are summarized by cognitive function composite in Table 17 and visually depicted in Figure 6. Overall, significant relationships were noted between at least one cognitive function composite and one or more polymorphisms of all candidate genes except *CMC2*, *MMP11*, and *RACGAP1*. Comprehensive results from the individual SNP and cognitive function regression analyses are

located in Table 18. Selected significant results, highlighting trends across cognitive function composites, are reported in the text to follow by candidate gene. All reported regression coefficients are from robust multiple linear regression models and convey the magnitude and direction of possession of one or more minor alleles on cognitive function performance.

Significant relations were observed between one or more polymorphisms in *ESR1* and performance for every cognitive function composite evaluated. Both main and interaction effects were observed. Poorer cognitive function performance was associated with possession of one or more minor alleles for the following *ESR1* SNPs regardless of group membership: *ESR1*rs2347867 and attention (b=-0.248, p=0.040), mental flexibility (b=-0.317, p=0.001), and psychomotor speed (b=-0.227, p=0.016); *ESR1*rs6557171 and attention (b=-0.234, p=0.047) and mental flexibility (b=-0.273, p=0.003); and *ESR1*rs7761846 and executive function (b=-0.388, p=0.005) and visual working memory (b=-0.351, 0.022). In contrast, better cognitive function performance was associated with possession of one or more *ESR1*rs488133 minor alleles and executive function (p=0.207, b=0.015) and psychomotor speed (b=0.275, p=0.004) scores, regardless of cohort membership. Significant interactions were also observed for *ESR1*rs488133; specifically, the combination of possession of one or more minor *ESR1*rs488133 alleles and prescribed anastrozole only (AO) group membership contributed positively to scores for concentration [SNP (b=-0.327, p=0.058); SNP-by-AO group (b=0.718, p=0.003)] and negatively to all three memory cognitive function composites, including verbal [SNP (b=0.365, p=0.018); SNP-by-AO group (b=-0.425, p=0.051)], visual [SNP (b=0.241, p=0.012); SNP-by-AO group (b=-0.288, p=0.034)], and visual working [SNP (b=0.516, 0.001); SNP-by-AO group (b=-0.544, 0.011)] memory. Prescribed chemotherapy plus anastrozole (CA) group membership interactions for all three memory cognitive function composites were also observed; in

particular; the combination of possession of one or more *ESR1*rs2941740 minor alleles and prescribed chemotherapy plus anastrozole group membership contributed positively to verbal [SNP (b=-0.318, p=0.041), SNP-by-CA group (b=0.663, p=0.008)], visual [SNP (b=-0.033, p=0.726), SNP-by-CA group (b=0.294, p=0.053)], and visual working [SNP (b=-0.258, p=0.105), SNP-by-CA group (b=0.554, p=0.030)] memory performance.

Associations were also reported between polymorphisms and pretreatment cognitive function in the gene neighboring *ESR1*, *CCDC170*. While significant relationships were also observed between *CCDC170* polymorphisms and executive function, verbal memory, and visual memory, the most noteworthy findings were related to SNP main effects for concentration. Possession of one or more minor alleles contributed negatively to concentration performance for four out of the five *CCDC170* SNPs examined regardless of study cohort membership: *CCDC170*rs12662670 (b=-0.363, p=0.016), *CCDC170*rs3734805 (b=-0.304, p=0.042), *CCDC170*rs3757318 (b=-0.334, p=0.042), and *CCDC170*rs6929137 (b=-0.322, p=0.003).

Significant findings related to polymorphisms in the other hormone receptor gene evaluated as part of this study, *PGR*, and concentration, executive function, psychomotor speed, verbal memory, visual memory, and visual working memory were reported as well. Numerous observed interaction effects indicate that variability in the *PGR* gene is particularly important to executive function performance in women with breast cancer. While possession of one or more *PGR*rs1042838 (b=0.337, p=0.025), *PGR*rs474320 (b=0.341, p=0.022), *PGR*rs484389 (b=0.318, p=0.024), or *PGR*rs608995 (b=0.324, p=0.020) minor alleles contribute positively to executive function performance, SNP-by-prescribed treatment group interaction effects for *PGR*rs1042838 [SNP-by-CA group (b=-0.155, p=0.528), SNP-by-AO group (b=-0.323, p=0.151)], *PGR*rs474320 [SNP-by-CA group (b=-0.392, p=0.116), SNP-by-AO group (b=-0.321,

p=0.168)], *PGR*rs484389 [SNP-by-CA group (b=-0.529, p=0.016), SNP-by-AO group (b=-0.494, p=0.015)], and *PGR*rs608995 [SNP-by-CA group (b=-0.538, p=0.013), SNP-by-AO group (b=-0.464, p=0.019)] contribute negatively to models counteracting the main effects and contributing an overall negative input to executive function performance in multiple instances.

Notable SNP main effects and/or SNP-by-prescribed treatment group interaction effects were observed for all *CCNB1*polymorphisms evaluated as part of this study. Significant SNP main effects were reported for *CCNB1*rs164390 and *CCNB1*rs350099, with possession of one or more minor alleles contributing to poorer concentration (b=-0.230, p=0.035; b=-0.246, p=0.026, respectively) performance regardless of study cohort membership. Whereas possession of one or more *CCNB1*rs164390 or *CCNB1*rs350099 minor alleles alone contributed positively to executive function (b=0.280, p=0.039; b=0.236, p=0.087, respectively), verbal memory (b=0.299, p=0.052; b=0.350, p=0.024, respectively), and visual working memory (b=0.581, p<0.001; b=0.562, p<0.001, respectively) scores, SNP-by-prescribed treatment interaction effects contribute negatively to executive function [SNP-by-CA group (b=-0.478, p=0.026), SNP-by-AO group (b=-0.551, p=0.006); SNP-by-CA group (b=-0.345, p=0.110), SNP-by-AO group (b=-0.507, p=0.011), respectively] , verbal memory [SNP-by-CA group (b=-0.355, p=0.143), SNP-by-AO group (b=-0.518, p=0.021); SNP-by-CA group (b=-0.445, p=0.065), SNP-by-AO group (b=-0.685, p=0.002), respectively], and visual working memory [SNP-by-CA group (b=-0.549, p=0.025), SNP-by-AO group (b=-0.540, p=0.017); SNP-by-CA group (b=-0.553, p=0.024), SNP-by-AO group (b=-0.608, p=0.008), respectively] scores. This same trend is observed for visual memory. In contrast, possession of one or more *CCNB1*rs350104 minor alleles alone contributed negatively to executive function (b=-0.225, p=0.177), verbal memory (b=-0.441, p=0.016), visual memory (b=-0.153, p=0.176), and visual working memory (b=-

0.475, p=0.009) scores, while SNP-by-prescribed treatment interaction effects contributed positively to executive function [SNP-by-CA group (b=0.128, p=0.606), SNP-by-AO group (b=0.470, p=0.032)], verbal memory [SNP-by-CA group (b=0.564, p=0.039), SNP-by-AO group (b=0.516, p=0.032)], visual memory [SNP-by-CA group (b=0.193, p=0.251), SNP-by-AO group (b=0.297, p=0.045)], and visual working memory [SNP-by-CA group (b=0.631, p=0.020), SNP-by-AO group (b=0.699, p=0.004)] scores.

Variability in *GSTM1* impacted similar cognitive function composites as *CCNB1*. Possession of one or more *GSTM1*rs412543 minor alleles contributed negatively to executive function (b=-0.291, p=0.009), verbal memory (b=-0.470, p=0.017), visual memory (b=-0.188, p=0.013), and visual working memory (b=-0.901, p<0.001) performance; SNP-by-prescribed treatment interaction effects contributed positively to verbal memory [SNP-by-CA group (b=0.624, p=0.043), SNP-by-AO group (b=0.640, p=0.034)] and visual working memory [SNP-by-CA group (b=1.012, p=0.001), SNP-by-AO group (b=0.473, p=0.116)] performance as well.

Variation in *BCL2* was also associated with performance for a number of cognitive function composites, including concentration, executive function, mental flexibility, psychomotor speed, verbal memory, and visual memory. Possession of one or more *BCL2*rs1564483 minor alleles contributed positively to executive function (b=0.191, p=0.030), mental flexibility (b=0.396, p=0.017), and visual memory (b=0.226, p=0.031) performance; significant SNP-by-prescribed anastrozole only group membership effects were also observed for mental flexibility (b=-0.607, p=0.015) and visual memory (b=-0.386, p=0.015).

Likewise, variation in *DIAPH3* was associated with performance for multiple cognitive function composites, including concentration, executive function, mental flexibility, verbal memory, visual memory, and visual working memory. Significant *DIAPH3*rs1337652 SNP main

effects were reported for executive function (b=-0.295, p=0.036) and visual memory (b=-0.126, p=0.032), whereas significant *DIAPH3*rs4547237 SNP main effects were reported for all three memory composites, including verbal (b=0.198, p=0.030), visual (b=0.194, p=0.039), and visual working (b=0.195, p=0.034) memory. Significant interactions were also observed between *DIAPH3*rs1337652 and mental flexibility [SNP (b=0.362, p=0.021), SNP-by-CA group (b=-0.361, p=0.166), SNP-by-AO group (b=-0.620, p=0.004)] and visual working memory [SNP (b=0.095, p=0.550), SNP-by-CA group (b=0.002, p=0.995), SNP-by-AO group (b=-0.446, p=0.042)] as well as *DIAPH3*rs4547237 and concentration [SNP (b=-0.273, p=0.118), SNP-by-CA group (b=0.053, p=0.847), SNP-by-AO group (b=0.610, p=0.014)] and executive function [SNP (b=0.410, p=0.002), SNP-by-CA group (b=-0.280, p=0.184), SNP-by-AO group (b=-0.389, p=0.040)].

Possession of one or more *MYBL2*rs2070235 minor alleles contributed positively to scores for attention (b=0.311, p=0.045) and negatively to verbal (b=-0.356, p=0.003), visual (b=-0.156, p=0.035), and visual working memory (b=-0.411, p=0.001) performance regardless of study cohort membership. Possession of one or more minor *MYBL2*rs619289 alleles also contributed negatively to verbal (b=-0.214, p=0.031) and visual working (b=-0.222, p=0.028) memory performance regardless of cohort membership. A significant *MYBL2*rs11556379 main effect for mental flexibility performance and significant *MYBL2*rs11556379 interaction effects for both prescribed treatment groups with executive function performance were also noted.

In addition, *MELK*rs10973007 SNP-by-prescribed treatment group interaction effects were reported for concentration [SNP (b=0.080, p=0.659), SNP-by-CA group (b=-0.112, p=0.701), SNP-by-AO group (b=-0.504, p=0.046)] and executive function [SNP (b=-0.215, p=0.138), SNP-by-CA group (b=0.601, p=0.010), SNP-by-AO group (b=0.462, p=0.022)], and

*MELK*rs2250340 SNP-by-prescribed treatment group interaction effects were reported for executive function [SNP (b=-0.069, p=0.671), SNP-by-CA group (b=0.935, p=0.005), SNP-by-AO group (b=-0.182, p=0.512)] and visual working memory [SNP (b=-0.003, p=0.988), SNP-by-CA group (b=0.604, p=0.101), SNP-by-AO group (b=-0.765, p=0.014)]. Possession of one or more *BAG1*rs706118 minor alleles was associated with poorer executive function (b=-0.312, p=0.027), visual memory (b=-0.216, p=0.025), and visual working memory (b=-0.310, p=0.001) scores, and possession of one or more *MKI67*rs1073248 minor alleles was associated with poorer mental flexibility (b=-0.264, p=0.007) and psychomotor speed (b=-0.194, p=0.047) performance. Likewise, possession of one or more *SCUBE2*rs6486125 minor alleles contributed negatively to executive function (b=-0.320, p=0.020) and mental flexibility (b=-0.193, p=0.043) performance scores regardless of study cohort membership; the combination of possession of one or more *SCUBE2*rs6486125 minor alleles and prescribed chemotherapy plus anastrozole group membership contributed negatively to attention (b=-0.653, p=0.043) performance scores.

Possession of one or more *MIR125A* (which impacts expression of *ERBB2*) rs12976445 minor alleles negatively impacted mental flexibility (b=-0.223, p=0.017) performance regardless of study cohort membership; a significant interaction effect between possession of *MIR125A*rs12976445 minor alleles and chemotherapy plus anastrozole group membership and attention [SNP (b=-0.619, p=0.001), SNP-by-CA group (b=0.643, p=0.032)] was also noted. Significant SNP-by-prescribed chemotherapy plus anastrozole only group membership effects were also observed for *BIRC5*rs1508147 and *BIRC5*rs9904341 and visual working memory (b=0.472, p=0.048; b=-0.510, p=0.041, respectively) performance.

*NDC80*rs12408485 and *NDC80*rs2292274 SNP-by-prescribed anastrozole only group membership was significantly associated with mental flexibility (b=0.589, p=0.010; b=-0.447,

p=0.046, respectively) performance. Significant SNP-by-prescribed anastrozole only group membership effects were also reported for *CD68*rs9901673 and verbal [SNP (b=0.446, p=0.010), SNP-by-AO group (b=-0.555, p=0.020)] and visual working [SNP (b=0.412, p=0.019), SNP-by-AO group (b=-0.498, p=0.040)] memory. Significant interactions were observed between *CENPA*rs3806517 and concentration [SNP (b=-0.363, p=0.039), SNP-by-CA group (b=0.056, p=0.849), SNP-by-AO group (b=0.615, p=0.012)], and *CENPA*rs3806518 and psychomotor speed [SNP (b=-0.148, p=0.335), SNP-by-CA group (b=0.079, p=0.745), SNP-by-AO group (b=0.507, p=0.018)] and verbal memory [SNP (b=0.282, p=0.063), SNP-by-CA group (b=-0.501, p=0.037), SNP-by-AO group (b=-0.423, p=0.045)] as well.

3.5.4 GRSs and cognitive function

One or more polymorphisms from the following genes, through either main SNP effects or SNP-by-prescribed treatment group interaction effects, were included in cognitive function composite GRSs: Attention – *ERBB2*(*MIR125A*), *ESR1*, *MYBL2*, and *SCUBE2*; Concentration – *AURKA*, *BCL2*, *CCNB1*, *CENPA*, *DIAPH3*, *ESR1*, *ESR1*(*CCDC170*), *GRB7*, *MELK*, and *PGR*; Executive Function – *BAG1*, *BCL2*, *CCNB1*, *CTSL2*, *DIAPH3*, *ESR1*, *GSTM1*, *MELK*, *MYBL2*, *PGR*, and *SCUBE2*; Mental Flexibility – *BCL2*, *DIAPH3*, *ERBB2*(*MIR125A*), *ESR1*, *GSTM1*(*NFE2L2*), *MKI67*, *NDC80*, *RFC4*, *RRM2*, and *SCUBE2*; Psychomotor Speed – *BCL2*, *CENPA*, *ESR1*, *MKI67*, and *PGR*; Verbal Memory – *AURKA*, *BCL2*, *CCNB1*, *CD68*, *CENPA*, *CTSL2*, *DIAPH3*, *ESR1*, *ESR1*(*CCDC170*), *GSTM1*, *MYBL2*, *NDC80*, *ORC6*, and *PGR*; Visual Memory – *BAG1*, *BCL2*, *CCNB1*, *DIAPH3*, *ESR1*, *GSTM1*, *MYBL2*, *PGR*, and *RRM2*; and Visual Working Memory – *AURKA*, *BAG1*, *BIRC5*, *CCNB1*, *CD68*, *DIAPH3*, *ESR1*, *GRB7*, *GSTM1*, *MELK*, *MYBL2*, and *PGR*. All GRSs were found to be significantly ($p<0.001$) related to the respective

cognitive function composite score (Table 17). Reported associations were all positive such that as GRS increases, cognitive function performance score increases as well (Figure 7).

3.5.5 *ERBB2* and *MKI67* subset analysis

Results for the *ERBB2* by HER2 IHC classification score genetic subset analysis are reported in Table 19. Two significant SNP-by-HER2 IHC classification score interactions for visual working memory performance were observed with *ERBB2*rs1058808 (b=-0.292, p=0.033) and *ERBB2*rs4252596 (b=-0.278, p=0.043).

Results for the *MKI67* by Ki67 index genetic subset analysis are reported in Table 20. One significant interaction between possession of one or more *MKI67*rs10732438 minor alleles and Ki67 index and mental flexibility was observed [Ki67 index (b=-0.015, p=0.006), SNP (b=-0.621, p=0.004), SNP-by-Ki67 index (b=0.016, p=0.023)].

3.6 RESULTS - CLINICOPATHOLOGIC TUMOR FEATURE AND CANDIDATE GENE INTERACTION ANALYSIS

The only significant ($p<0.05$) interaction between CTFs and GRSs was related to the visual working memory composite and observed between tumor location octant, specifically lower outer compared to upper outer, and visual working memory GRS (Table 21). A list of all significant ($p<0.05$) individual CTFs tested for interactions with GRSs by cognitive function composite is presented in Table 22.

In order to further evaluate the potential ability of combining CTF and genetic information to account for a greater amount of observed variance in cognitive function performance, regression models containing marginally significant ($p<0.20$) CTF predictors and GRSs were developed. CTFs included in the final model for each cognitive function composite were identified from robust regression models that incorporated all statistically significant individual and interaction terms (Table 23). Table 24 presents the R^2 for the final CTF models as well as the R^2 for the final combined CTF plus GRS models for each cognitive function composite. These results must be interpreted with extreme caution due to exploratory model building and very small sample sizes.

3.7 DISCUSSION

In this study investigating the impact of variation in CTFs of breast cancer and host DNA variation in genes used to clinically evaluate the biology of breast cancer on cognitive performance in postmenopausal women with early-stage breast cancer, we found evidence to support the hypothesis that heterogeneity in molecular and pathologic characteristics of breast cancer account for variability in pretreatment cognitive function performance.

Overall, we found that CTFs related to cancer stage, tumor size, tumor focality, tumor location, histologic type and grade, hormone receptor and HER2 expression, cellular proliferation, as well as Oncotype DX® Breast Cancer Assay Recurrence Score® and Magee Equation recurrence score were significantly ($p<0.05$) associated with performance for one or more cognitive function composites. However, the most intriguing findings were related to memory performance and HER2 status or HER2 IHC classification score. For all memory

composites, possession of a HER2 positive tumor contributed to poorer performance compared to possession of a HER2 negative tumor. Likewise, as HER2 IHC classification scores increased, memory performance scores decreased. HER2 is an epidermal growth factor receptor encoded by the *ERBB2* (*erb-b2 receptor tyrosine kinase 2*) gene. Within the context of breast cancer, we commonly discuss the oncogenic role of amplification of HER2 and its use as an indicator of poorer breast cancer prognosis (Iqbal & Iqbal, 2014; Moasser, 2007; Wesoła & Jeleń, 2015). Our findings suggest that poorer breast cancer prognosis, based on HER2 expression, is associated with poorer pretreatment memory performance. These cognition-related findings are further strengthened when we consider the important and widespread proto-oncogenic role that *ERBB2* plays in proper neural development (Britsch et al., 1998; Kim, Sun, Oglesbee, & Yoon, 2003; Kornblum, Yanni, Easterday, & Seroogy, 2000; Lee et al., 1995; Thompson et al., 2007).

Complementary genetic analyses of *ERRB2* polymorphisms were also conducted. Bearing in mind the role of *ERRB2* in brain development, we were initially surprised to find significant cognitive performance associations with only one polymorphism, *MIR125A (microRNA 125a*) rs12976445, which impacts expression levels of *ERBB2*; preliminary evidence suggests that this polymorphism may have merit as a prognostic biomarker of survival in women with breast cancer (Jiao et al., 2014). However, because *ERBB2* has rare or limited expression in normal breast tissue (Pavelic et al., 1992; Stark et al., 2000), we theorized that polymorphisms may only be impactful in women with breast cancer with HER2 amplifications. We tested this hypothesis in the subset of women with breast cancer who provided genetic specimens. The sub-analysis revealed two significant interaction effects for visual working memory performance between rs1058808(Pro1170Ala, C>G) in the 3' untranslated region of *ERBB2*, and rs2517955,

an upstream intronic variant, and HER2 IHC classification score. As HER2 classification scores increased, the magnitude of the negative contribution to performance score also increased in women with breast cancer with rs1058808-GC+CC or rs2517955-TC+CC genotypes, lending support to our conjecture. Interestingly, rs1058808 genotype has been associated with breast tumor HER2 expression; Su and colleagues reported that women with breast cancer with rs1058808-CG+GG genotypes were more likely to have higher HER2 expression levels than women with breast cancer with the rs1058808-CC genotype (2015). Additionally, higher rates of Parkinson's disease have been reported in female rs1058808 Ala allele (CG+GG genotypes) carriers (Wang et al., 2013).

One other interesting finding from the CTF analysis that warrants further discussion was the impact of tumor location, specifically tumor location octant on cognitive functioning. We found that overall tumor location octant was related to pretreatment mental flexibility performance. Most significantly, women with a tumor in the lower inner octant displayed poorer pretreatment mental flexibility performance compared to women with a tumor in the upper outer octant. Depending on quadrant classification, previous studies have reported associations between the lower, inner, and lower inner breast quadrants and inferior outcomes, including decreased survival and disease recurrence (Arriagada et al., 2002; Colleoni et al., 2005; Hazrah, Dhir, Gupta, Deo, & Parshad, 2009; Kamakura, Akazawa, Nomura, Sugimachi, & Nose, 1996; Lohrisch, Jackson, Jones, Mates, & Olivotto, 2007; Sarp et al., 2007); although, conflicting results have also been reported (Wu et al., 2014). In addition, upper outer quadrant location, the most common location for breast tumors, has been associated with better prognosis compared to other tumor locations (Bao, Yu, Jiang, Shao, & Di, 2014; Sohn, Arthurs, Sebesta, & Brown, 2008). These differences in outcomes are hypothesized to occur because of undetected breast

cancer spread to the internal mammary lymph nodes (Estourgie, Nieweg, Olmos, Rutgers, & Kroon, 2004; Sarp et al., 2007; Sohn et al., 2008). While different classifications limit interpretation of results, it is fascinating to note that the tumor location most strongly related to poorer mental flexibility performance has also been associated with poorer breast cancer outcomes.

In order to further facilitate interpretation of the tumor location results, we evaluated how CTFs differed by tumor location octant. This analysis aided in the interpretation of the finding that women with breast cancer in the lower middle octant had poorer visual working memory performance when compared to women with tumors in the upper outer octant. While multiple CTFs varied by octant, two differences stood out. A higher percentage of lower middle octant tumors were HER2 positive compared to the other tumor locations, and the lower middle octant displayed higher mean HER2 IHC classification scores. These differences, which relate back to previously discussed associations between memory and HER2 amplification, suggest that CTFs overrepresented in a particular octant may be driving relationships between location and cognitive function rather than the actual location itself. While one published expert opinion suggested that HER2 expression does not vary by anatomic location within the breast (Jasani, Novelli, Ruschoff, & Osamura, 2010), no formal studies have been conducted.

Shifting attention towards the results of the genetic analysis, we reported significant relationships between performance on at least one cognitive function composite and one or more polymorphisms of all candidate genes evaluated, with the exception of *CMC2*, *MMP11*, and *RACGAP1*, by either SNP main effects (i.e., observed variability in cognitive function performance in women with breast cancer and healthy controls is associated with a certain polymorphism) and/or SNP-by-prescribed treatment group interaction effects (i.e., observed

variability in cognitive function performance is associated with a certain combination of polymorphism and prescribed treatment group). Significant findings related to the candidate genes found to most broadly impact cognitive function performance across multiple cognitive function composites, specifically, *ESR1* and *CCDC170*, *PGR*, *CCNB1*, *MYBL2*, *BCL2*, *GSTM1*, and *DIAPH3*, are discussed, as exemplars, in detail below.

Performance on every cognitive function composite was related to *ESR1* (*estrogen receptor 1*) polymorphisms either through main effects and/or interactions effects. These findings were not unanticipated due to previously reported relationships between *ESR1* polymorphisms and cognitive outcomes, including functioning, impairment, and Alzheimer's disease (Sundermann, Maki, & Bishop, 2011). The most global associations with a single *ESR1* polymorphism in the current study occurred with an intronic upstream variant, rs488133. The effects of this polymorphism on cognitive function performance were different by cognitive composite and study cohort: rs488133-CT+TT contributed positively to executive function and psychomotor speed performance in all study participants. rs488133-CT+TT negatively impacted concentration performance in healthy controls, but positively impacted concentration performance in women with breast cancer prescribed anastrozole only. In contrast, rs488133-CT+TT positively impacted memory performance in healthy controls, but negatively impacted memory performance in women with breast cancer prescribed anastrozole only. In addition, while reported in other investigations of middle-aged and older women, we did not observe global cognitive impairment trends or memory deficits related to two well-studied polymorphisms in exon 1 of *ESR1* named for the respective restriction enzyme recognition sites, PvuII (rs2234693) and Xbal (rs9340799) (Bousman et al., 2012; Kravitz, Meyer, Seeman, Greendale, & Sowers, 2006; Yaffe et al., 2009; Yaffe, Lui, Grady, Stone, & Morin, 2002).

Polymorphisms in *CCDC170* (*coiled-coil domain containing 170*), the upstream neighbor of *ESR1*, were included in this study to more fully represent variability in *ESR1*. Associations between *CCDC170* polymorphisms and breast cancer susceptibility, progression, and survival have been reported (Fletcher et al., 2011; Hein et al., 2012; Hong et al., 2014; Stevens et al., 2011; Yamamoto-Ibusuki et al., 2015). In addition, *ESR1-CCDC170* chromosomal rearrangements have been associated with more aggressive estrogen receptor positive breast cancers (Veeraraghavan et al., 2015). While the function of *CCDC170* is unknown and no studies to date have investigated associations between *CCDC170* polymorphisms and cognitive phenotypes, results from this analysis, in which possession of one or more *CCDC170* minor alleles in four (rs12662670, rs3734805, rs3757318, rs6929137) out of the five SNPs evaluated was related to poorer concentration performance in all study participants, suggest that variation in *CCDC170* plays an important role in concentration.

The physiologic effects of progesterone are mediated by the protein encoded by *PGR* (*progesterone receptor*). Progesterone receptors are expressed throughout the brain in every neural cell type (Brinton et al., 2008). Henderson et al. found that progesterone concentrations were significantly and positively related to global cognition and verbal memory performance in healthy women less than 6 years since menopause (2013). Moreover, Voytko, Murray, and Higgs found that estrogen plus progesterone improved executive function and attention performance in surgically menopausal monkeys (2009). For executive function performance, we observed significant interactions between multiple *PGR* polymorphisms and study cohort. In all instances, possession of *PGR*rs1042838-GT+TT, *PGR*rs474320-TA+AA, *PGR*rs484389-TC+CC, or *PGR*rs608995-AT+TT genotypes contributed positively to executive function performance scores in healthy controls. When we looked at the interaction of these minor alleles

within the context of breast cancer, we saw the opposite effect; the combination of possession of one or more minor alleles and membership in a breast cancer cohort was found to negatively impact scores, offsetting the positive SNP main effects and contributing an overall negative input to executive function performance in multiple instances. The first SNP, rs1042838(Val660Leu, G>T), is a missense polymorphism (i.e., alters the amino acid sequence of a protein) in exon 4 that is in linkage disequilibrium with a silent polymorphism in exon 5, rs1042839(His770His, C>T), and a 320bp Alu element insertion at intron G; this polymorphic variant is collectively known as PROGINS. While the functional consequences remain unclear, variant PROGINS has been associated with increased breast and ovarian cancer risk (Agoulnik et al., 2004; Liu et al., 2014; Rockwell et al., 2012; Romano, Delvoux, Fischer, & Groothuis, 2007; Stenzig et al., 2012). rs474320 is an intronic variant reported to be in tight linkage with PROGINS (Lee et al., 2010). rs1042839 was also evaluated in this study and, as expected, generated very similar results to rs1042838; discrepancies in call rate may account for the differences in significance. rs484389 and rs608995 are located in the 3' untranslated region of *PGR*. Taken together, these findings indicate that variation in regulation of progesterone receptors may affect executive function performance and, furthermore, that the polymorphic impact on performance may vary in the systemic environment of a healthy individual compared to that of an individual diagnosed with breast cancer.

CCNB1 (c*yclin B1*) encodes a cell cycle regulatory protein important in mitosis (NCBI Resource Coordinators, 2015). Because expression levels from this gene are used in three out of five of the prognostic multigene expression profiles for breast cancer from which candidate genes were identified, *CCNB1* was one of our top candidates for investigation of study hypotheses. Significant interactions were reported with study cohort for three functional

polymorphisms, rs164390(102G>T), rs350099(-957C>T), and rs350104(-457C>T), located in the promotor region of *CCNB1* and memory and executive function performance. In general, we found that possession of rs164390GT+TT or rs350099CT+CC genotypes contributed positively to performance scores in healthy controls but close to zero or negatively in women with breast cancer. The opposite contribution was observed for rs350104CT+CC genotypes. The genotypes associated with poorer cognitive performance in the cohorts of women with breast cancer, rs164390-GT+TT, rs350099-CT+CC, and rs350104-TT, are all hypothesized to lead to lower levels of *CCNB1*expression via reduced recruitment of transcription factors to the promotor region of the gene (Silvestre-Roig et al., 2014). This result is contradictory to anticipated findings as higher cyclin B levels in breast tissue are associated with more severe cancer phenotypes (Kawamoto, Koizumi, & Uchikoshi, 1997; Winters et al., 2001). In addition, cyclin B levels were reported to be upregulated in autopsy hippocampal tissue in individuals with neuropathological Alzheimer's disease and clinical dementia compared to individuals with normal aging (Silva et al., 2014). Nevertheless, the consistency of findings across three variants all theorized to impact expression in the same direction, lends support to these associations. We would like to point of that one or more polymorphisms in the four other genes represented in three prognostic multigene expression profiles for breast cancer, *CENPA*, *MELK*, *MYBL2*, and *ORC6*, were associated with performance on at least one cognitive function composite.

MYBL2 (*MYB proto-oncogene like 2*) encodes a nuclear protein, B-MYB, involved in cell cycle progression and promotion of cell survival through activation of anti-apoptotic genes (NCBI Resource Coordinators, 2015; Sala, 2005). However, overexpression of B-MYB in certain settings induces apoptosis and has been reported to contribute to neuronal cell death (Iyirhiaro et al., 2014; Liu, Biswas, & Greene, 2004; Lui, Nath, Chellappan, & Greene, 2005;

Sala, 2005). We found significant relationships with two missense polymorphisms in *MYBL2*, rs11556379(Ile624Met, C>G) and rs2070235(Ser427Gly, A>G). The minor alleles of these polymorphisms have been reported to alter protein conformation, impair regulation of downstream targets, decrease anti-apoptotic activity, and reduce cancer risk (Schwab et al., 2008). Interestingly, for all study participants rs2070235-AG+GG genotypes contributed positively to attention and negatively to memory performance scores, while rs11556379-CG+GG genotypes contributed positively to mental flexibility performance scores. We also reported a significant interaction related to executive function where rs11556379-CG+GG genotypes had the opposite impact on performance in healthy controls (positive contribution to scores) and women with breast cancer (negative contribution to scores). Additionally, we reported associations between polymorphisms in a gene regulated by *MYBL2* that is also involved in apoptosis, *BCL2* (*B-cell CLL/lymphoma 2*), and concentration, executive function, mental flexibility, psychomotor speed, verbal memory, and visual memory performance. *BCL2* expression has been associated with prognostication of disease free survival, overall survival, and recurrence in breast cancer (Abdel-Fatah et al., 2010; Aleskandarany et al., 2015; Bremer et al., 2009; Callagy, Webber, Pharoah, & Caldas, 2008; Dawson et al., 2010; Kerr & Wittliff, 2011; Linke, 2006; Lyng et al., 2013). Moreover, normal breast tissue from women with breast cancer was reported to display higher levels of *BCL2* expression than breast tissue from women with no evidence of cancer (Batchelder, Gordon-Weeks, & Walker, 2009). In relation to neurologic phenotypes, polymorphisms in *BCL2* have been found to impact outcomes after traumatic brain injury (Hoh et al., 2010) and have been associated with hippocampal volume (Sloan et al., 2010).

One of the functional polymorphisms located in the promoter region of *GSTM1* (*glutathione S-transferase mu 1*), rs412543(-498C>G), was found to be important for memory and executive function performance. *GSTM1* encodes an enzyme with antioxidant properties that detoxifies electrophilic compounds, including carcinogens, drugs, and environmental toxins, throughout the body (NCBI Resource Coordinators, 2015). By decreasing the binding capability of the transcription factor AP-2 to the *GSTM1* promotor region, the G allele has been reported to decrease *GSTM1* transcription by 30-40% compared to the C allele (Yu et al., 2009). Both decreased and enhanced (attributed to counterproductive depletion of glutathione) *GSTM1* expression has been associated with increased breast cancer risk (Reed, 1990; Roodi, Dupont, Moore, & Parl, 2004; Yu et al., 2009). We found that rs412543-GG+CG, and hypothesized decreased *GSTM1* expression, contributed negatively to executive function and memory performance in all study participants. However, we also found positive interaction effects between rs412543-GG+CG and breast cancer cohort related to verbal and visual working memory. While the mechanism is unclear, the paradoxical quality of *GSTM1* under and over expression combined with study results suggests that decreased or moderate *GSTM1* expression may be beneficial to certain aspects of cognitive function in women with breast cancer. Considering the detoxification properties of *GSTM1*, further evaluation of cognitive decline over time in women with breast cancer receiving adjuvant chemotherapy and/or antiestrogen therapy is recommended.

Variation in the two upstream intronic polymorphisms selected to represent *DIAPH3* (*diaphanous related formin 3*), rs1337652 and rs4547237, were associated with performance for multiple cognitive composites as well. *DIAPH3* is involved in actin remodeling and regulation of cell movement and adhesion (NCBI Resource Coordinators, 2015). *DIAPH3* downregulation

and silencing has been associated with metastatic disease due to loss of normal gene function and acquisition of an amoeboid cancer cell phenotype (Hager et al., 2012). Evidence also suggests that *DIAPH3* is critical to brain development and is involved in cell migration, the formation of dendrites and axons, axon guidance, and synaptic activity (Vorstman et al., 2011).

Each of the candidate genes discussed previously with the most significant findings from our analysis are represented by multiple functional and/or tagging SNPs and are well-described in the literature. This is not the case for the three candidate genes with no reported associations with pretreatment cognitive performance. Single SNPs, rs131451 and rs7303531, were included in the analysis for *MMP11* (*matrix metallopeptidase 11*) and *RACGAP1* (*Rac GTPase activating protein 1*), respectively. Both SNPs are upstream variants. No associations have been reported between *MMP11* or *RACGAP1* and cognitive phenotypes in the literature. *CMC2* (*C-x(9)-C motif containing 2*) is an even more poorly described and studied gene with reported involvement in cytochrome c oxidase activity (Horn et al., 2010). Two upstream (rs1025065 and rs1981867) polymorphisms and one downstream (rs9936489) polymorphism were identified using the Phase III HapMap database based on National Center for Biotechnology Information gene location (Chr16: 80975802..81006897) as *CMC2* is not a displayed gene in HapMap. We must be mindful that our analysis is limited to the information known about these genes and the polymorphisms described at the current time.

Because of the complexity of breast cancer as a disease and cognitive function as a phenotype, we calculated weighted GRSs for each cognitive function composite to evaluate the collective effect of possession of multiple risk or protective minor alleles of genes used to clinically evaluate the biology of breast cancer. Every GRS was significantly ($p<0.001$) and positively associated with its respective cognitive function composite. When the GRSs were

added as predictors to regression models, including age, estimated verbal intelligence, levels of depressive symptoms, anxiety, fatigue, and pain, and prescribed treatment group, the explained variance (R^2) increased by 0.066 to 0.244 for each cognitive function composite. This substantial increased in R^2 speaks to both the importance of host variation in genes used to clinically evaluate the biology of breast cancer to pretreatment cognitive function performance as well as the use of multiple common variants, plus personal and environmental factors, to model a complex phenotype.

Unfortunately, due to limitations in sample size, results from the interaction analysis as well as the full pathologic and genetic predictive model analysis have limited interpretability. Small sample sizes also limited our ability to conduct genetic analyses by genotype rather than by the presence or absence or one or more minor alleles; thus, we were unable to evaluate gene-dosage effects. Other limitations to this study were also noted. As with any retrospective chart review, CTF data were limited to availability in the medical record and recommended testing at the time of diagnosis (e.g., lack of Ki67 proliferative marker evaluation in participants enrolled at the beginning of the parent study). The study sample was comprised of postmenopausal women with hormone receptor positive, early-stage breast cancer who were primarily Caucasian; the generalizability of study findings to premenopausal women, hormone negative, in situ and more advanced breast cancers, or more diverse populations is unknown. Additionally, analyses were not adjusted for the potentially confounding effects of surgery and/or anesthesia exposure on pretreatment cognitive performance.

We would also like to acknowledge this study's many strengths, including: 1) hypothesis driven aims; 2) biologically-plausible candidate gene selection; 3) evaluation of both functional and tagging polymorphisms to fully represent genetic variability in candidate genes; 4) well-

characterized cognitive function phenotype; 5) adjustment for potential covariates/confounders of cognitive function, including age, intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain; 6) consideration of breast cancer heterogeneity through use of prescribed treatment group and inclusion of a matched control group in the candidate gene analysis; 7) adjustment for prescribed treatment group to account for noted differences in mean cognitive performance unrelated to genotype in the candidate gene analysis; and 8) evaluation of the collective effect of multiple risk or protection polymorphisms using weighted GRSs.

While relationships between host DNA and cognitive function performance are advantageous for a number of reasons, including lack of change over time and tissue nonspecificity, associations with gene expression and protein levels warrant investigation as the most notable findings from this study were related to polymorphisms with known functional consequences or located in regulatory regions. We postulate that cognitive function performance variability in women with breast cancer may be at least partially driven by tumor gene expression and corresponding protein levels. Longitudinal studies that include cognitive assessment prior to primary surgery would be ideal for evaluation of the effect of tumor gene expression as well as changes in gene expression due to tumor removal and treatment of primary and secondary cancer sites on variability in cognitive function performance. Future analyses should also investigate the effect of variation in CTFs of breast cancer, host DNA in genes used to clinically evaluate the biology of breast cancer, and tumor expression levels on cognitive function throughout and following adjuvant therapy.

3.8 CONCLUSION

In summary, the objective of this study was to explore the hypothesis that heterogeneity in the biology of breast cancers is associated with variability in the presence and/or severity of pretreatment cognitive performance among postmenopausal women diagnosed with early-stage breast cancer. Significant associations between variation in CTFs, host polymorphisms from candidate genes used to clinically evaluate the biology of breast cancer, and computed GRSs and variability in pretreatment cognitive function performance support our hypothesis and suggest that inherent, disease-related features of breast cancer play a critical role in cognitive dysfunction symptomatology. These findings merit future investigation to further elucidate pathophysiologic mechanisms of and identify clinically relevant biomarkers for breast cancer related cognitive dysfunction.

APPENDIX A

PROPOSAL & PRELIMINARY WORK: TABLES AND FIGURES

Table 6: Proposed SNPs for Candidate Gene Analysis

Gene	Number of SNPs	SNP	Functional (F) or Tagging (T)
AURKA	8	rs1047972	F
		rs16979877	F
		rs2273535	F
		rs6024836	T
		rs6064389	F
		rs8173	F
		rs911161	F
		rs911162	F
BAG1	1	rs706118	F
BCL2	18	rs12961976	T
		rs1564483	F
		rs17759659	T
		rs1800477	F
		rs2279115	F
		rs4941195	T
		rs4987768	T
		rs4987852	T
		rs4987853	F
		rs4987855	T
		rs4987856	T
		rs7230177	T
		rs7231901	T
		rs7243091	T
		rs956572	F
		rs9807663	T
		rs10138824 (*MPP5*)	F
		rs8008724 (*EIF2S1*)	F
BIRC5	8	rs1042489	F
		rs1508147	T
		rs17878467	F
		rs2239680	F
		rs3764383	F
		rs8073069	F
		rs8073903	T
		rs9904341	F
CCNB1	4	rs164390	F
		rs350099	F
		rs350104	F
		rs352626	T
CD68	2	rs12942088	T
		rs8066665	T
CENPA	2	rs3806517	T
		rs3806518	T
CMC2	3	rs1025065	F
		rs1981867	T
		rs9936489	T
CTSL2	2	rs16919034	T
		rs4361859	T $_{(MAF=0.137)}$
DIAPH3	2	rs1337652	T

		rs4547237	T
ERBB2	13	rs1058808	F
		rs1136201	F
		rs1476278	T
		rs1810132	F
		rs2517955	T
		rs4252596	F
		rs4252633	F
		rs903501	T
		rs9303274	T
		rs61764370 (*KRAS*)	F
		rs12976445 (*MIR125A*)	F
		rs41275794 (*MIR125A*)	F
		rs491951 (*ZNF345*)	F
ESR1	56	rs10484919	F
		rs10484922	F
		rs1048919	T
		rs1062577	F
		rs11964281	F
		rs12173570	F
		rs12665044	F
		rs1514348	F
		rs1801132	F
		rs1884051	F
		rs2046210	F
		rs2071454	T
		rs2077647	F
		rs2228480	F
		rs2234693	F
		rs2347867	F
		rs2744677	F
		rs2813543	T
		rs2813544	T
		rs2941740	T
		rs3020314	F
		rs33778609	F
		rs34535804	T
		rs3778099	F
		rs3778609	F
		rs3798577	F
		rs488133	T
		rs532010	T
		rs6557171	F
		rs728524	F
		rs77275268	F
		rs7739506	T
		rs7761133	F
		rs7761846	F
		rs7766585	F
		rs7767143	T
		rs8179176	F
		rs827421	F
		rs851967	T

		rs851971	T
		rs851982	F
		rs851998	T
		rs910416	F
		rs9322331	F
		rs9340799	F
		rs9383938	F
		rs9383939	T
		rs9397435	F
		rs9397456	F
		rs9478245	T
		rs985694	F
		rs1038304 (*CCDC170*)	F
		rs12662670 (*CCDC170*)	F
		rs3734805 (*CCDC170*)	F
		rs3757318 (*CCDC170*)	F
		rs6929137 (*CCDC170*)	F
GRB7	1	rs9910678	T (MAF=0.035)
GSTM1	5	rs1065411	F
		rs142484086	F
		rs412543	F
		rs35652124 (*NFE2L2*)	F
		rs6721961 (*NFE2L2*)	F
MELK	3	rs10973007	T (MAF=0.177)
		rs2250340	T (MAF=0.164)
		rs3780350	T (MAF=0.164)
MKI67	2	rs10732438	T
		rs10764751	T
MMP11	2	rs131451	T (MAF=0.066)
		rs28382576	T
MYBL2	5	rs11556379	F
		rs2070235	F
		rs619289	F
		rs826943	F
		rs826944	F
NDC80	2	rs12408485	T
		rs2136241	T
ORC6	1	rs33994299	T
PGR	15	rs1042838	F
		rs1042839	F
		rs10895054	F
		rs10895068	F
		rs11224561	T
		rs1893505	T
		rs1942836	T
		rs471767	T
		rs474320	T
		rs4754732	T
		rs484389	T
		rs499590	T
		rs568157	T
		rs590688	F
		rs608995	T

RACGAP1	1	rs7303531	T $_{(MAF=0.058)}$
RFC4	1	rs1354091	T
RRM2	3	rs1138729	F
		rs4309551	T
		rs6759180	F
SCUBE2	3	rs1136966	T
		rs4910440	T
		rs6486125	T
Note. MAF=Minor Allele Frequency; SNP=Single Nucleotide Polymorphism. All tagging SNP MAFs ≥0.20 unless noted otherwise.			

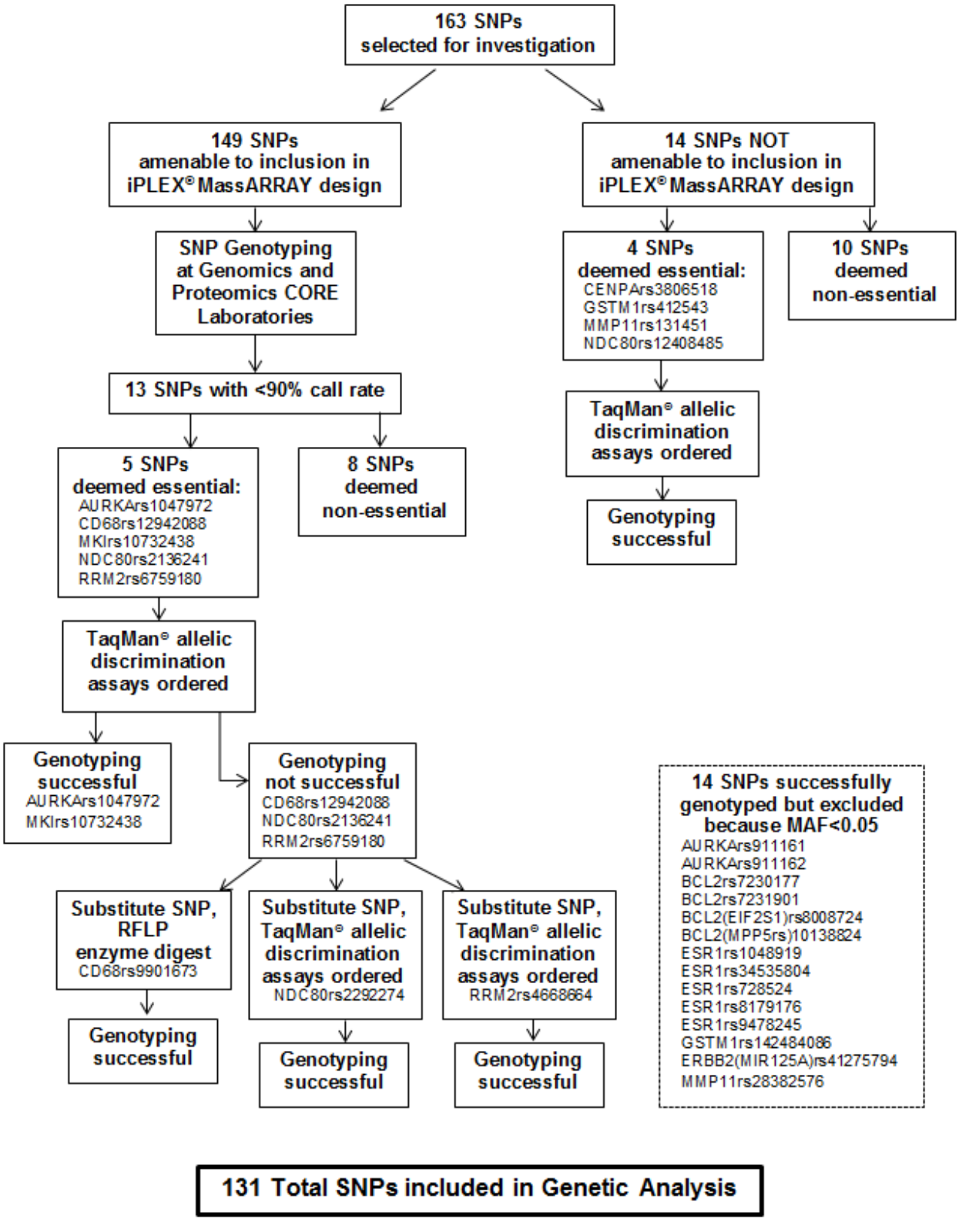


Figure 3: Dissertation SNP Genotyping Workflow

APPENDIX B

DATA-BASED MANUSCRIPT: TABLES AND FIGURES

Table 7: CTF Analysis Participant Characteristics (N=329)

Characteristic (Measure)	Mean±SD or n (%)	Minimum	Maximum
Age (years)	61.05±5.976	45	75
Education (years)	14.80±2.805	6	26
Estimated Verbal Intelligence (NART-R)	108.45±8.584	77.08	125.14
Depressive Symptoms (BDI-II)	5.33±5.619	0	32
Anxiety (POMS Tension-Anxiety subscale)	7.64±5.801	0	29
Fatigue (POMS Fatigue-Inertia subscale)	5.72±5.986	0	27
Pain (BPI)	1.44±2.165	0	9
Marital Status, currently married or living with significant other	223 (67.8)	-	-
Number of Children	1.89±1.237	0	7
Race, Caucasian	317 (96.4)	-	-
Cognitive Function Composite Z-Scores			
Attention, n=321	-0.1587±0.93945	-4.25	1.63
Concentration, n=328	-0.0141±0.91255	-3.41	3.98
Executive Function, n=329	-0.3953±0.63810	-2.37	1.83
Mental Flexibility, n=328	0.1197±0.78899	-4.05	1.63
Psychomotor Speed, n=329	-0.1201±0.92513	-6.01	2.28
Verbal Memory, n=329	-0.2088±0.66864	-2.58	1.28
Visual Memory, n=329	0.0680±0.66866	-3.28	0.86
Visual Working Memory, n=329	-0.0035±0.78009	-4.73	1.55

Note. BDI-II=Beck Depression Inventory-II; BPI=Brief Pain Inventory; CTF=Clinicopathologic Tumor Feature; NART-R=National Adult Reading Test-Revised; POMS=Profile of Mood States; SD=Standard Deviation. Only participants with complete confounder/covariate information are included in the participant characteristic statistics.

Table 8: CTF Summary Statistics (N=329)

Feature	Mean±SD or n (%)	Minimum	Maximum
Overall TNM Stage, n=329			
Stage 1	214 (65)	-	-
Stage 2a	75 (22.8)	-	-
Stage 2b	24 (7.3)	-	-
Stage 3a	16 (4.9)	-	-
Tumor Size (cm), n=328	1.66±1.500	0.10	14.00
Aggregate Tumor Size (cm), n=328	1.80±1.599	0.10	14.00
Tumor Stage, n=329			
T1a	37 (11.2)	-	-
T1b	82 (24.9)	-	-
T1c	133 (40.4)	-	-
T2	65 (19.8)	-	-
T3	12 (3.6)	-	-
Lymph Node, n=325			
Positive	73 (22.5)	-	-
Negative	252 (77.5)	-	-
Number of Positive Nodes, n=329	0.42±1.054	0	8
Tumor Focality, n=329			
Single	277 (84.2)	-	-
Multiple	52 (15.8)	-	-
Tumor Laterality, n=329			
Right Breast	149 (45.3)	-	-
Left Breast	180 (54.7)	-	-
Tumor Location Octant, n=321			
Upper Outer	130 (40.5)	-	-
Lower Outer	25 (7.8)	-	-
Upper Inner	40 (12.5)	-	-
Lower Inner	20 (6.2)	-	-
Upper Middle	41 (12.8)	-	-
Outer Middle	31 (9.7)	-	-
Lower Middle	20 (6.2)	-	-
Inner Middle	14 (4.4)	-	-
Invasive Type, n=329			
Ductal	285 (86.9)	-	-
Lobular	35 (10.7)	-	-
Ductal & Lobular	8 (2.4)	-	-
Nottingham Score, n=315	6.04±1.306	3	9
Nottingham Grade, n=316			
Grade 1	95 (30.1)	-	-
Grade 2	171 (54.1)	-	-
Grade 3	50 (15.8)	-	-
ER Status, n=328			
Positive	324 (98.8)	-	-
Negative	4 (1.2)	-	-
ER H-Score, n=311	256.90±59.978	0	300
Oncotype DX ER Score, n=102	10.287±1.056	7.8	12.5
PR Status, n=328			
Positive	288 (87.8)	-	-
Negative	40 (12.2)	-	-
PR H-Score, n=310	130.08±101.301	0	300
Oncotype DX PR Score, n=102	7.08±1.569	3.2	10.2

HER2 Status, n=318			
Positive	28 (8.8)	-	-
Negative	290 (91.2)	-	-
HER2 IHC Score, n=291	1.21±0.869	0	3
Oncotype DX HER2 Score, n=74	8.93±0.812	7.6	12.8
LV Invasion, n=323			
Present	68 (21.1)	-	-
Absent	255 (78.9)	-	-
Ki67 Classification, n=169			
Low	66 (39.1)	-	-
Moderate	50 (29.6)	-	-
High	34 (20.1)	-	-
Very High	19 (11.2)	-	-
Ki67 Index, n=168	23.10±21.522	1	90
Oncotype DX Recurrence Score®, n=160	18.26±9.76	0	63
Magee Equation Recurrence Score, n=298	20.51±7.77	1.92	48.87

Note. CTF=Clinicopathologic Tumor Feature; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; LV=Lymphovascular; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; SD=Standard Deviation; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors. Only participants with complete confounder/covariate information are included in the summary statistics.

Table 9: CTF and Cognitive Function Robust Regression Results

Regression coefficient, p-value	Attention	Concentration	Executive Function	Mental Flexibility	Psychomotor Speed	Verbal Memory	Visual Memory	Visual Working Memory
Overall TNM Stage	n=321	n=328	n=329	n=328	n=329	n=329	n=329	n=329
Stage 2a	0.116, 0.305	-0.065, 0.550	0.005, 0.949	0.064, 0.472	0.154, 0.075[#]	0.059, 0.485	0.033, 0.550	0.190, 0.033*
Stage 2b	-0.200, 0.262	0.139, 0.425	-0.146, 0.279	.008, 0.953	-0.211, 0.128[#]	-0.157, 0.244	-0.079, 0.363	-0.021, 0.885
Stage 3a	-0.143, 0.520	-0.207, 0.325	0.034, 0.833	0.179, 0.296	0.087, 0.601	0.068, 0.676	-0.181, 0.088[#]	-0.203, 0.239
Overall p-value	0.354	0.544	0.726	0.694	0.082[#]	0.503	0.218	0.076[#]
Ref: Stage 1								
Tumor Size (cm)	n=320	n=327	n=328	n=327	n=328	n=328	n=328	n=328
	-0.031, 0.318	-0.032, 0.283	-0.001, 0.955	0.009, 0.717	0.001, 0.951	0.018, 0.418	-0.003, 0.848	0.009, 0.715
Aggregate Tumor Size (cm)	n=320	n=327	n=328	n=327	n=328	n=328	n=328	n=328
	-0.027, 0.353	-0.032, 0.249	-0.004, 0.849	0.012, 0.600	-0.010, 0.653	0.004, 0.847	-0.009, 0.544	-0.002, 0.929
Tumor Stage	n=321	n=328	n=329	n=328	n=329	n=329	n=329	n=329
T1b	-0.023, 0.890	0.032, 0.842	0.021, 0.866	0.023, 0.858	0.018, 0.893	-0.132, 0.288	0.031, 0.710	0.042, 0.755
T1c	-0.021, 0.896	-0.093. 0.537	0.075, 0.522	0.180, 0.142[#]	0.035,0.774	0.068, 0.557	-0.010, 0.894	0.080, 0.521
T2	-0.013, 0.940	-0.032, 0.848	0.022, 0.868	0.100, 0.464	-0.007, 0.960	-0.051, 0.693	-0.003, 0.970	0.106. 0.446
T3	-0.248, 0.389	-0.238, 0.375	0.051, 0.806	0.133, 0.543	0.115, 0.598	0.140, 0.501	-0.009, 0.946	0.084. 0.708
Overall p-value	0.932	0.723	0.951	0.404	0.977	0.179	0.969	0.943
Ref: T1a								
Node Positive	n=318	n=324	n-325	n=324	n=325	n=325	n=325	n=325
Ref: Negative	0.058, 0.602	-0.077, 0.473	0.045, 0.587	0.144, 0.101[#]	0.093, 0.280	0.054, 0.514	0.010, 0.859	0.048, 0.595
Number of Positive Nodes	n=321	n=328	n=329	n=328	n=329	n=329	n=329	n=329
	0.018, 0.679	-0.024, 0.574	0.004, 0.892	0.057, 0.102[#]	0.040, 0.242	0.008, 0.806	-0.030, 0.172[#]	-0.003, 0.930
Multifocal	n=321	n=328	n=329	n=328	n=329	n=329	n=329	n=329
Ref: Single	-0.206, 0.098[#]	-0.056, 0.645	-0.071, 0.450	-0.073, 0.461	-0.092, 0.345	-0.278, 0.003*	-0.084, 0.170[#]	-0.108, 0.280
Left Breast	n=321	n=328	n=329	n=328	n=329	n=329	n=329	n=329
Ref: Right Breast	0.059, 0.521	-0.109, 0.226	0.019, 0.785	-0.092, 0.204	0.064, 0.366	0.156, 0.025*	-0.034, 0.452	0.163, 0.026*
Location Octant	n=314	n=321	n=322	n=321	n=322	n=322	n=322	n=322
Lower Outer	0.318, 0.090[#]	-0.083, 0.642	0.102, 0.456	-0.207, 0.145[#]	-0.122, 0.390	-0.140, 0.308	0.010, 0.912	0.180, 0.211

Upper Inner Lower Inner Upper Middle Outer Middle Lower Middle Inner Middle Overall p-value *Ref: Upper Outer*	-0.315, 0.040* -0.281, 0.167[#] -0.087, 0.555 0.065, 0.701 -0.158, 0.429 0.190. 0.415 0.092[#]	0.102, 0.490 0.167, 0.396 0.101, 0.490 -0.055, 0.740 -0.053, 0.787 -0.059, 0.804 0.933	-0.018, 0.874 -0.114, 0.446 0.098, 0.380 0.202, 0.109[#] 0.201, 0.182[#] 0.244, 0.165[#] 0.393	-0.171, 0.147[#] -0.511. 0.001* -0.180, 0.121[#] 0.058, 0.660 -0.258, 0.100[#] -0.008, 0.967 0.027*	-0.088, 0.455 -0.263, 0.092[#] -0.109, 0.347 -0.063, 0.631 0.109, 0.483 -0.111, 0.542 0.647	0.035, 0.762 -0.028, 0.851 0.020, 0.862 0.050, 0.693 -0.198, 0.191[#] -0.104, 0.553 0.813	-0.054, 0.466 0.056, 0.569 0.092, 0.209 0.069, 0.401 -0.088, 0.370 0.095, 0.410 0.612	-0.180, 0.133[#] -0.165, 0.297 0.022, 853 -0.012, 0.927 -0.432, 0.007* 0.120, 0.515 0.052[#]
Invasive Type Lobular Ductal & Lobular Overall p-value *Ref: Ductal*	n=320 -0.080, 0.595 0.505, 0.085[#] 0.186[#]	n=327 -0.043, 0.772 -0.088, 0.760 0.920	n=328 0.090, 0.432 -0.215, 0.335 0.441	n=327 -0.062, 0.603 -0.219, 0.349 0.579	n=328 -0.034, 0.771 -0.209, 0.361 0.640	n=328 0.107, 0.346 -0.003, 0.990 0.640	n=328 -0.072, 0.335 -0.090, 0.537 0.536	n=328 -0.054, 0.656 0.303, 0.203 0.388
Nottingham Score	n=308 -0.032, 0.381	n=314 -0.002, 0.958	n=315 0.029, 0.291	n=314 -0.018, 0.542	n=315 0.005, 0.844	n=315 0.045, 0.093[#]	n=315 0.011, 0.523	n=315 0.048, 0.100[#]
Nottingham Grade Grade 2 Grade 3 Overall p-value *Ref: Grade 1*	n=309 0.077, 0.467 -0.124, 0.392 0.312	n=315 -0.033, 0.751 0.048, 0.733 0.817	n=316 0.085, 0.288 0.132, 0.235 0.417	n=315 0.000, 0.995 -0.070. 0.554 0.799	n=316 0.066, 0.417 -0.009, 0.940 0.632	n=316 0.111, 0.162[#] 0.051, 0.641 0.368	n=316 0.037, 0.484 0.035, 0.635 0.771	n=316 0.084, 0.329 0.202, 0.090[#] 0.231
ER Positive *Ref: Negative*	n=320 0.298, 0.472	n=327 -0.290, 0.473	n=328 0.211, 0.498	n=327 0.391, 0.233	n=328 -0.231, 0.471	n=328 0.022, 0.945	n=328 -0.047, 0.819	n=328 -0.269, 0.420
ER H-Score	n=303 0.001, 0.165[#]	n=310 -0.001, 0.409	n=311 0.001, 0.397	n=310 0.001, 0.301	n=311 -0.001, 0.229	n=311 0.000, 0.730	n=311 0.000, 0.606	n=311 0.000, 0.679
Oncotype DX ER Score	n=100 -0.010, 0.904	n=102 -0.004, 0.960	n=102 0.082, 0.195[#]	n=102 0.049, 0.393	n=102 0.014, 0.836	n=102 -0.013, 0.842	n=102 0.074, 0.069[#]	n=102 0.047, 0.488
PR Positive *Ref: Negative*	n=320 -0.129, 0.364	n=327 -0.004, 0.975	n=328 -0.140, 0.184[#]	n=327 0.024, 0.834	n=328 0.069, 0.524	n=328 -0.256, 0.015*	n=328 -0.050, 0.470	n=328 -0.048, 0.669
PR H-Score	n=302 0.000, 0.668	n=309 0.000, 0.405	n=310 0.000, 0.967	n=309 0.000, 0.461	n=310 0.000, 0.923	n=310 0.000, 0.806	n=310 0.000, 0.538	n=310 0.000, 0.655
Oncotype DX	n=100	n=102	n=102	n=102	n=102	n=102	n=102	n=102

PR Score	-0.062, 0.275	0.046. 0.405	-0.029, 0.498	0.010, 0.794	0.031, 0.506	-0.019, 0.650	0.007, 0.782	-0.057, 0.196[#]
HER2 Positive *Ref: Negative*	n=310 -0.185, 0.283	n=317 0.277, 0.087[#]	n=318 -0.131, 0.293	n=317 -0.065, 0.626	n=318 -0.031, 0.803	n=318 -0.287, 0.018*	n=318 -0.270, 0.001*	n=318 -0.490, <0.001*
HER2 IHC Classification Score	n=285 -0.031, 0.591	n=290 0.001, 0.979	n=291 -0.018, 0.672	n=290 -0.019, 0.679	n=291 -0.038, 0.389	n=291 -0.072, 0.093[#]	n=291 -0.081, 0.003*	n=291 -0.170, <0.001*
Oncotype DX HER2 Score	n=73 -0.106, 0.388	n=74 0.105, 0.365	n=74 -0.008, 0.925	n=74 0.024, 0.753	n=74 -0.051, 0.627	n=74 -0.041, 0.670	n=74 0.011, 0.862	n=74 -0.099, 0.288
LV Invasion *Ref: No Invasion*	n=315 -0.094, 0.420	n=322 -0.138, 0.216	n=323 0.078, 0.368	n=322 0.102, 0.264	n=323 -0.076, 0.394	n=323 0.095, 0.270	n=323 0.018, 0.753	n=323 0.059, 0.524
Ki67 Classification Moderate High Very High Overall p-value *Ref: Low*	n=168 0.001, 0.995 0.093, 0.589 0.140, 0.507 0.872	n=168 0.381, 0.009* 0.035, 0.829 0.254, 0.205 0.042*	n=169 -0.003, 0.983 -0.065, 0.636 0.287, 0.090[#] 0.266	n=168 -0.146, 0.243 0.068, 0.635 -0.059, 0.740 0.481	n=169 -0.064, 0.589 -0.193, 0.158[#] 0.122, 0.463 0.325	n=169 -0.086, 0.476 -0.139, 0.312 0.015, 0.927 0.711	n=169 -0.019, 0.799 0.044, 0.602 0.031, 0.765 0.894	n=169 -0.085, 0.475 -0.011, 0.936 0.073, 0.662 0.799
Ki67 Index	n=167 0.003, 0.359	n=167 0.003, 0.342	n=168 0.003, 0.135[#]	n=167 0.001, 0.712	n=168 0.002, 0.343	n=168 0.000, 0.906	n=168 0.001, 0.722	n=168 0.001, 0.540
Oncotype DX Recurrence Score®	n=157 0.000, 0.957	n=159 0.004, 0.483	n=160 -0.003, 0.582	n=160 -0.010, 0.032*	n=160 -0.005, 0.387	n=160 0.005, 0.367	n=160 -0.003, 0.388	n=160 0.009, 0.115[#]
Magee Equation Recurrence Score	n=291 -0.008, 0.238	n=297 0.005, 0.447	n=298 0.000, 0.918	n=297 -0.002, 0.701	n=298 0.002, 0.609	n=298 -0.002, 0.615	n=298 -0.002, 0.497	n=298 -0.002, 0.667

Note. *=p<0.05, [#]=0.05<p<0.20; CTF=Clinicopathologic Tumor Feature; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; LV=Lymphovascular; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors; Ref=Reference Group. All regression coefficient estimates and p-values reported from robust multiple linear regression models generated using Huber weighting and biweighting iterations. All models are adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 10: CTF Squared Continuous Predictor and Cognitive Function Robust Regression Results (p<0.20)

Regression coefficient, p-value	Attention	Concentration	Executive Function	Mental Flexibility	Psychomotor Speed	Verbal Memory	Visual Memory	Visual Working Memory
Tumor Size (cm) *Squared variable*	n=320 -0.050, 0.477 0.002, 0.775	n=327 -0.003, 0.964 -0.003, 0.656	n=328 -0.004, 0.946 0.000, 0.962	n=327 0.035, 0.528 -0.003, 0.618	n=328 0.003, 0.957 0.000, 0.977	n=328 0.036, 0.485 -0.002, 0.714	n=328 -0.021, 0.539 0.002, 0.560	n=328 0.081, 0.143# -0.009, 0.135#
Oncotype DX ER Score *Squared variable*	n=100 -2.390, 0.059# 0.117, 0.061#	n=102 0.085, 0.947 -0.004, 0.944	n=102 -0.985, 0.307 0.053, 0.269	n=102 -0.339, 0.695 0.019, 0.654	n=102 -0.761, 0.473 0.039, 0.462	n=102 -0.866, 0.364 0.042, 0.370	n=102 -0.514, 0.373 0.029, 0.312	n=102 -0.817, 0.421 0.043, 0.391
Oncotype DX PR Score *Squared variable*	n=100 -0.749, 0.081# 0.050, 0.105#	n=102 -0.096, 0.811 0.010, 0.718	n=102 -0.299, 0.327 0.020, 0.370	n=102 0.424, 0.118# -0.030, 0.126#	n=102 0.439, 0.186# -0.029, 0.218	n=102 0.209, 0.484 -0.017, 0.441	n=102 -0.147, 0.448 0.011, 0.418	n=102 -0.353, 0.264 0.021, 0.342
HER2 IHC Classification Score *Squared variable*	n=285 0.115, 0.475 -0.056, 0.339	n=290 -0.095, 0.549 0.036, 0.525	n=291 -0.023, 0.848 0.002, 0.966	n=290 0.054, 0.666 -0.029, 0.527	n=291 -0.094, 0.444 0.023, 0.597	n=291 0.048, 0.689 -0.046, 0.280	n=291 -0.025, 0.736 -0.022, 0.424	n=291 -0.009, 0.943 -0.061, 0.172#
Oncotype DX HER2 Score *Squared variable*	n=73 0.678, 0.592 -0.040, 0.539	n=74 -1.454, 0.195# 0.081, 0.165#	n=74 -0.732, 0.432 0.037, 0.440	n=74 -1.002, 0.208 0.053, 0.196#	n=74 -0.801, 0.464 0.039, 0.493	n=74 0.858, 0.388 -0.047, 0.366	n=74 1.983, <0.001* -0.104, <0.001*	n=74 0.037, 0.969 -0.007, 0.889
Ki67 Index *Squared variable*	n=167 0.003, 0.728 0.000, 0.938	n=167 0.000, 0.964 0.000, 0.807	n=168 0.000, 0.951 0.000, 0.578	n=167 0.003, 0.680 0.000, 0.749	n=168 -0.010, 0.206 0.000, 0.100#	n=168 -0.010, 0.183# 0.000, 0.155#	n=168 0.005, 0.264 0.000, 0.291	n=168 0.001, 0.881 0.000, 0.972
Oncotype DX Recurrence Score® *Squared variable*	n=157 -0.008, 0.665 0.000, 0.665	n=159 -0.021, 0.225 0.001, 0.127#	n=160 -0.014, 0.348 0.000, 0.423	n=160 -0.007, 0.603 0.000, 0.781	n=160 0.018, 0.228 -0.001, 0.086#	n=160 0.032, 0.021* -0.001, 0.029*	n=159 -0.004, 0.706 0.000, 0.896	n=159 -0.002, 0.927 0.000, 0.457
Magee Equation Recurrence Score *Squared variable*	n=291 -0.001, 0.981 0.000, 0.804	n=297 -0.019, 0.495 0.000, 0.386	n=298 0.012, 0.579 0.000. 0.588	n=297 0.009, 0.693 0.000, 0.621	n=298 0.000, 0.990 0.000, 0.894	n=298 0.019, 0.347 0.000, 0.283	n=298 0.026, 0.055# -0.001, 0.034*	n=298 0.051, 0.028* -0.001, 0.021*

Note. *=p<0.05, #=0.05<p<0.20; CTF=Clinicopathologic Tumor Feature; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor. All regression coefficient estimates and p-values reported from robust multiple linear regression models generated using Huber weighting and biweighting iterations. All models are adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 11: CTFs by Tumor Location Octant (p<0.20)

Mean±SD, n or n (%)	**Tumor Location Octant** (N=321)								F-test[a] or Fisher's Exact Test[b] p-value
	Upper Outer	**Lower Outer**	**Upper Inner**	**Lower Inner**	**Upper Middle**	**Lower Middle**	**Outer Middle**	**Inner Middle**	
Tumor Size (cm)	1.66±1.374, n=130	1.47±1.407, n=25	1.81±1.453, n=40	1.87±1.163, n=19	1.25±0.971, n=41	1.34±0.688, n=20	2.43±2.795, n=31	1.14±0.526, n=14	p=0.035
Aggregate Tumor Size (cm)	1.78±1.472, n=130	1.77±1.800, n=25	1.91±1.561, n=40	1.94±1.320, n=19	1.37±1.029, n=41	1.57±1.022, n=20	2.58±2.780, n=31	1.14±0.526, n=14	p=0.058
Tumor Focality									
Single	115 (88.5)	19 (76.0)	36 (90.0)	18 (90.0)	30 (73.2)	15 (75.0)	24 (77.4)	13 (92.9)	p=0.118
Multiple	15 (11.5)	6 (24.0)	4 (10.0)	2 (10.0)	11 (26.8)	5 (25.0)	7 (22.6)	1 (7.1)	
Invasive Type									
Ductal	115 (89.1)	21 (84.0)	34 (85.0)	18 (90.0)	39 (95.1)	18 (90.0)	20 (64.5)	14 (100.0)	p=0.041
Lobular	11 (8.5)	2 (8.0)	6 (15.0)	1 (5.0)	2 (4.9)	2 (10.0)	9 (29.0)	0 (0.0)	
Ductal & Lobular	3 (2.3)	2 (8.0)	0 (0.0)	1 (5.0)	0 (0.0)	0 (0.0)	2 (6.5)	0 (0.0)	
Nottingham Grade									
Grade 1	45 (35.7)	4 (17.4)	18 (46.2)	5 (25.0)	10 (26.3)	3 (15.8)	5 (16.1)	5 (38.5)	p=0.036
Grade 2	63 (50.0)	17 (73.9)	13 (33.3)	9 (45.0)	23 (60.5)	14 (73.7)	19 (61.3)	8 (61.5)	
Grade 3	18 (14.3)	2 (8.7)	8 (20.5)	6 (30.0)	5 (13.2)	2 (10.5)	7 (22.6)	0 (0.0)	
HER2 Status									
Negative	114 (89.8)	24 (96.0)	36 (94.7)	19 (100.0)	35 (89.7)	14 (70.0)	26 (92.9)	14 (100.0)	p=0.082
Positive	13 (10.2)	1 (4.0)	2 (5.3)	0 (0.0)	4 (10.3)	6 (30.0)	2 (7.1)	0 (0.0)	
HER2 IHC Score	1.21±0.882, n=119	0.86±0.640, n=22	1.39±0.688, n=36	1.18±0.728, n=17	1.22±0.886, n=37	1.94±0.929, n=16	1.21±1.021, n=24	0.67±0.778, n=12	p=0.003
LV Invasion									
Absent	96 (76.2)	24 (96.0)	29 (72.5)	17 (85.0)	26 (66.7)	15 (75.0)	26 (83.9)	14 (100.0)	p=0.035
Present	30 (23.8)	1 (4.0)	11 (27.5)	3 (15.0)	13 (33.3)	5 (25.0)	5 (16.1)	0 (0.0)	

Note. CTF=Clinicopathologic Tumor Feature; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; LV=Lymphovascular; SD=Standard Deviation. [a]One-way ANOVAs utilized to compare means of continuous variables. [b]Fisher's Exact Tests computed using 2-sided Monte Carlo sampling based on 10,000 sampled tables used to examine associations between categorical variables.

Table 12: Individual CTFs (p<0.20) Tested for Two-way Interactions

Cognitive Function Composite	CTF with p<0.20 in Model with Linear Term	CTF with p<0.20 in Model with Quadratic Term
Attention	• Tumor Focality • Tumor Location Octant • Invasive Type • ER H-Score	• Oncotype DX ER Score • Oncotype DX PR Score
Concentration	• HER2 Status (+/-) • Ki67 Classification	• Oncotype DX HER2 Score • Oncotype DX Recurrence Score®
Executive Function	• Tumor Location Octant • Oncotype DX ER Score • PR Status (+/-) • Ki67 Classification • Ki67 Index	
Mental Flexibility	• Tumor Stage • Node Status (+/-) • Number of Positive Nodes • Tumor Location Octant • Oncotype DX Recurrence Score®	• Oncotype DX PR Score
Psychomotor Speed	• Overall TNM Stage • Tumor Location Octant • Ki67 Classification	• Oncotype DX PR Score • Ki67 Index • Oncotype DX Recurrence Score®
Verbal Memory	• Tumor Focality • Tumor Laterality • Tumor Location Octant • Nottingham Score • Nottingham Grade • PR Status (+/-) • HER2 Status (+/-) • HER2 IHC Classification Score	• Ki67 Index • Oncotype DX Recurrence Score®
Visual Memory	• Overall TNM Stage • Number of Positive Nodes • Tumor Focality • Oncotype DX ER Score • HER2 Status (+/-) • HER2 IHC Classification Score	• Oncotype DX HER2 Score • Magee Equation Recurrence Score
Visual Working Memory	• Overall TNM Stage • Tumor Laterality • Tumor Location Octant • Nottingham Score • Nottingham Grade • Oncotype DX PR Score • HER2 Status (+/-) • HER2 IHC Classification Score • Oncotype DX Recurrence Score®	• Tumor Size • HER2 IHC Classification Score • Magee Equation Recurrence Score

Note. CTF=Clinicopathologic Tumor Feature; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; LV=Lymphovascular; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors.

Table 13: Significant Two-way CTF Interactions (p<0.05)

Cognitive Function Composite	Significant (p<0.05) Interaction Effect (Main Effect regression coefficient, p-value)	Interaction Effect regression coefficient, p-value
Attention		
	Lower Inner Octant (-0.388, 0.075) by	n=313
	Lobular Type (-0.303, 0.261)	1.814, 0.043
	ER H-Score (-0.138, 0.005) by	n=96
	Oncotype DX PR Score (-12.082, 0.004)	0.040, 0.006
	Oncotype DX PR Score Squared (0.834, 0.006)	-0.003, 0.009
Concentration		
	Oncotype DX HER2 Score (20.372, 0.145) by	n=73
	Oncotype DX Recurrence Score® (13.291, 0.041)	-3.030, 0.033
	Oncotype DX Recurrence Score® Squared (-0.376, 0.026)	0.086, 0.021
	Oncotype DX HER2 Score Squared (-1.148, 0.128) by	
	Oncotype DX Recurrence Score®	0.171, 0.026
	Oncotype DX Recurrence Score® Squared	-0.005, 0.017
Executive Function		
	Lower Inner Octant (4.417, 0.037) by	n=101
	Oncotype DX ER Score (0.157, 0.140)	-0.457, 0.027
	Upper Inner Octant (0.620, 0.012) by	n=166
	Ki67 Moderate Classification (0.146, 0.531)	-0.914, 0.031
	Upper Inner Octant (0.497, 0.030) by	n=167
	Ki67 Index (0.007, 0.100)	-0.017, 0.026
	Very High Ki67 Classification (7.789, 0.004) by	n=76
	Oncotype DX ER Score (0.187, 0.096)	-0.705, 0.011
	Ki67 Index (0.088, 0.005) by	n=76
	Oncotype DX ER Score (0.209, 0.031)	-0.008, 0.012
Mental Flexibility		
	Lower Inner Octant (-23.333, 0.026) by	n=101
	Oncotype DX PR Score (0.754, 0.011)	6.968, 0.020
	Oncotype DX PR Score Squared (-0.047, 0.023)	-0.511, 0.015
Psychomotor Speed		
	TNM Stage 3a (-0.267, 0.368) by	n=169
	Ki67 Moderate Classification (-0.190, 0.257)	1.034, 0.038
	Upper Inner Octant (-0.053, 0.830) by	n=166
	Very High Ki67 Classification (-0.049, 0.864)	-2.352, <0.001
	Outer Middle Octant (-0.698, 0.009) by	
	Moderate Ki67 Classification (-0.153, 0.407)	0.940, 0.032
	Outer Middle Octant (-0.698, 0.009) by	
	High Ki67 Classification (-0.301, 0.199)	1.086, 0.031
	Moderate Ki67 Classification (-7.160, 0.016) by	n=76
	Oncotype DX PR Score (-1.780, 0.018)	2.418, 0.006
	Oncotype DX PR Score Squared (0.140, 0.010)	-0.191, 0.003
	High Ki67 Classification (-9.843, 0.002) by	
	Oncotype DX PR Score	3.045, 0.002
	Oncotype DX PR Score Squared	-0.222, 0.001
	Very High Ki67 Classification (-26.790, 0.019) by	
	Oncotype DX PR Score	8.737, 0.013
	Oncotype DX PR Score Squared	-0.670, 0.012
	TNM Stage 3a (-1.353, 0.038) by	n=168
	Ki67 Index (-0.016, 0.133)	0.188, 0.019
	Ki67 Index Squared (0.000, 0.074)	-0.004, 0.022
	Upper Inner Octant (-0.358, 0.302) by	n=166

	Ki67 Index (-0.012, 0.337)	0.037, 0.123
	Ki67 Index Squared (0.000, 0.357)	-0.001, 0.004
Verbal Memory		
	PR Positive (-0.375, 0.001) by	n=318
	HER2 Positive (-0.785, 0.002)	0.611, 0.033
	Lower Inner Octant (-2.558, 0.023) by	n=157
	Oncotype DX Recurrence Score® (0.014, 0.557)	0.282, 0.015
	Oncotype DX Recurrence Score® Squared (0.000, 0.974)	-0.007, 0.017
Visual Memory		
	Oncotype DX ER Score (0.090, 0.042) by	n=99
	HER2 Positive (6.353, 0.001)	-0.720, <0.001
	TNM Stage 3a (-3.725, 0.010) by	n=298
	Magee Equation Recurrence Score (-0.036, 0.049)	0.249, 0.024
	Magee Equation Recurrence Score Squared (-0.001, 0.029)	-0.004, 0.047
Visual Working Memory		
	Lower Outer Octant (-0.343, 0.132) by	n=321
	Left Breast (0.106, 0.344)	0.780, 0.007
	Lower Outer Octant (0.175, 0.227) by	n=311
	HER2 Positive (-0.267, 0.162)	-1.627, 0.018
	Outer Middle Octant (0.659, 0.038) by	n=158
	Oncotype DX Recurrence Score® (0.009, 0.295)	-0.036, 0.010
	Nottingham Score (0.170, 0.001) by	n=278
	HER2 IHC Classification Score (0.368, 0.106)	-0.085, 0.015
	Nottingham Grade 3 (0.722, 0.001) by	n=279
	HER2 IHC Classification Score (0.039, 0.703)	-0.392, 0.006
	HER2 IHC Classification Score (0.150, 0.268) by	n=142
	Oncotype DX Recurrence Score® (0.032, 0.005)	-0.013, 0.025
	HER2 IHC Classification Score (-0.352, <0.001) by	n=290
	Tumor Size (-0.194, 0.170)	0.163, 0.037
	Tumor Size Squared (0.027, 0.255)	-0.020, 0.109
	Lower Outer Octant (-0.291, 0.333) by	n=283
	HER2 IHC Classification Score (-0.183, 0.365)	1.502, 0.016
	HER2 IHC Classification Score Squared (0.030, 0.678)	-0.819, 0.008
	Oncotype DX Recurrence Score® (0.008, 0.578) by	n=142
	HER2 IHC Classification Score (-0.648, 0.164)	0.033, 0.123
	HER2 IHC Classification Score Squared (0.264, 0.138)	-0.014, 0.031

Note. CTF=Clinicopathologic Tumor Feature; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors. All regression coefficient estimates and p-values reported from robust multiple linear regression models generated using Huber weighting and biweighting iterations. The reference group for tumor location octant is the upper outer octant. The reference for invasive type is ductal. The reference group for Ki67 classification is low. The reference group for TNM Stage is Stage 1. The reference group for PR status is PR negative. The reference group for HER2 status is HER2 negative. The reference group for the left breast is the right breast. The reference group for Nottingham Grade is Grade 1. All models are adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 14: Genetic Analysis Overall Participant Characteristics (N=220)

Characteristic (Measure)	Mean±SD or n (%)	Minimum	Maximum
Age (years)	60.02±6.086	43	75
Education (years)	15.12±2.970	9	29
Estimated Verbal Intelligence (NART-R)	110.39±9.113	82.42	127.81
Depressive Symptoms (BDI-II)	4.83±4.957	0	29
Anxiety (POMS Tension-Anxiety subscale)	7.47±5.536	0	27
Fatigue (POMS Fatigue-Inertia subscale)	5.61±5.942	0	27
Pain (BPI)	1.30±2.126	0	9
Marital Status, currently married or living with significant other	139 (63.2)	-	-
Number of Children	2.00±1.403	0	8
Race, Caucasian	209 (95.0)	-	-
Cognitive Function Composite Z-Scores			
Attention, n=219	-0.1070±0.94939	-4.02	1.70
Concentration, n=219	-0.0560±0.83170	-2.20	2.50
Executive Function, n=220	-0.2357±0.64539	-1.69	2.41
Mental Flexibility, n=219	0.0965±0.75203	-3.64	1.73
Psychomotor Speed, n=220	-0.0548±0.88616	-3.67	1.22
Verbal Memory, n=220	-0.1087±0.72263	-1.77	1.67
Visual Memory, n=220	0.0832±0.68602	-4.63	0.86
Visual Working Memory, n=220	0.0358±0.77624	-3.02	1.33

Note. BDI-II=Beck Depression Inventory-II; BPI=Brief Pain Inventory; NART-R=National Adult Reading Test-Revised; POMS=Profile of Mood States; SD=Standard Deviation. Only participants with complete confounder/covariate information are included in the participant characteristic statistics.

Table 15: Genetic Analysis Participant and Tumor Characteristics by Study Cohort

Characteristic (Measure) Mean±SD or n (%)	Prescribed CA n=55	Prescribed AO n=83	Healthy Controls n=82	[a]F-test or Chi-Square/ Fisher's Exact Test[b] p-value
Age (years)	58.76±5.467	62.47±5.964	58.39±5.858	p<0.001*
Education (years)	15.67±2.783	14.95±3.056	14.93±2.993	p=0.285
Estimated Verbal Intelligence (NART-R)	108.94±8.871	107.04±8.844	114.74±7.796	p<0.001*
Depressive Symptoms (BDI-II)	5.24±4.615	4.60±4.650	4.79±5.495	p=0.757
Anxiety (POMS Tension-Anxiety subscale)	9.61±6.140	6.97±4.654	6.55±5.619	p=0.003*
Fatigue (POMS Fatigue-Inertia subscale)	5.11±5.329	5.84±6.352	5.72±5.955	p=0.763
Pain (BPI)	1.47±1.961	1.55±2.265	0.93±2.059	p=0.130
Marital Status, currently married or living with significant other	38 (69.1)	54 (65.1)	47 (57.3)	p=0.348
Number of Children	1.75±1.220	2.05±1.387	2.13±1.522	p=0.266
Race, Caucasian	52 (94.5)	81 (97.6)	76 (92.7)	p=0.305
Cognitive Function Composite Z-Scores				
Attention, n=219	-0.052±0.937	-0.202±1.017	-0.047±0.889	p=0.513
Concentration, n=219	-0.204±0.667	-0.010±0.904	-0.005±0.850	p=0.322
Executive Function, n=220	-0.218±0.599	-0.490±0.509	0.010±0.705	p<0.001*
Mental Flexibility, n=219	0.164±0.656	0.090±0.786	0.055±0.783	p=0.707
Psychomotor Speed, n=220	0.071±0.845	-0.240±0.954	0.048±0.819	p=0.054
Verbal Memory, n=220	0.018±0.662	-0.341±0.638	0.041±0.786	p=0.001*
Visual Memory, n=220	0.287±0.352	0.009±0.708	0.022±0.803	p=0.038*
Visual Working Memory, n=220	0.299±0.514	-0.064±0.741	-0.039±0.913	p=0.014*
Overall TNM Stage, n=130				
Stage 1	22 (44.0)	65 (81.3)	-	p<0.001*
Stage 2a	17 (34.0)	13 (16.3)	-	
Stage 2b	6 (12.0)	2 (2.5)	-	
Stage 3a	5 (10.0)	0 (0)	-	
Tumor Size (cm), n=129	2.16±1.484	1.23±0.709	-	p<0.001*
Aggregate Tumor Size (cm), n=130	2.35±1.591	1.39±0.964	-	p<0.001*
Tumor Stage, n=130				
T1a	1 (2.0)	13 (16.3)	-	p=0.001*
T1b	8 (16.0)	23 (28.7)	-	
T1c	21 (42.0)	33 (41.3)	-	
T2	16 (32.0)	11 (13.8)	-	
T3	4 (8.0)	0 (0)	-	
Lymph Node, n=129				
Positive	19 (38.0)	5 (6.3)	-	p<0.001*
Negative	31 (62.0)	74 (93.7)	-	
Number of Positive Nodes, n=130	0.94±1.789	0.06±0.244	-	p<0.001*
Tumor Focality, n=130				
Single	41 (82.0)	63 (78.8)	-	p=0.652
Multiple	9 (18.0)	17 (21.3)	-	
Tumor Laterality, n=130				
Right Breast	21 (42.0)	36 (45.0)	-	p=0.737
Left Breast	29 (58.0)	44 (55.0)	-	
Tumor Location Octant, n=125				
Upper Outer	21 (43.8)	29 (37.7)	-	p=0.982
Lower Outer	2 (4.2)	6 (7.8)	-	
Upper Inner	6 (12.5)	9 (11.7)	-	
Lower Inner	2 (4.2)	5 (6.5)	-	

Upper Middle	8 (16.7)	10 (13.0)	-	
Outer Middle	4 (8.3)	8 (10.4)	-	
Lower Middle	2 (4.2)	4 (5.2)	-	
Inner Middle	3 (6.3)	6 (7.8)	-	
Invasive Type, n=129				
Ductal	45 (90.0)	63 (79.7)	-	p=0.323
Lobular	5 (10.0)	14 (17.7)	-	
Ductal & Lobular	0 (0)	2 (2.5)	-	
Nottingham Score, n=125	6.60±1.370	5.72±1.122	-	p<0.001*
Nottingham Grade, n=125				
Grade 1	9 (18.0)	27 (36.0)	-	p<0.001*
Grade 2	26 (52.0)	44 (58.7)	-	
Grade 3	15 (30.0)	4 (5.3)	-	
ER Status, n=130				
Positive	48 (96.0)	80 (100)	-	p=0.146
Negative	2 (4.0)	0 (0)	-	
ER H-Score, n=124	240.08±73.684	265.87±44.592	-	p=0.017*
Oncotype DX ER Score, n=46	9.79±1.138	10.41±0.934	-	p=0.053
PR Status, n=130				
Positive	38 (76.0)	71 (88.8)	-	p=0.055
Negative	12 (24.0)	9 (11.3)	-	
PR H-Score, n=124	110.35±101.612	129.69±97.208	-	p=0.289
Oncotype DX PR Score, n=46	6.22±1.591	7.31±1.391	-	p=0.020*
HER2 Status, n=125				
Positive	9 (19.1)	4 (5.1)	-	p=0.017*
Negative	38 (80.9)	74 (94.9)	-	
HER2 IHC Score, n=120	1.40±0.984	1.10±0.858	-	p=0.081
Oncotype DX HER2 Score, n=33	8.76±1.079	8.87±0.396	-	p=0.657
LV Invasion, n=127				
Present	21 (42.9)	6 (7.7)	-	p<0.001*
Absent	28 (57.1)	72 (92.3)	-	
Ki67 Classification, n=68				
Low	10 (38.5)	18 (42.9)	-	p=0.114
Moderate	5 (19.2)	14 (33.3)	-	
High	6 (23.1)	9 (21.4)	-	
Very High	5 (19.2)	1 (2.4)	-	
Ki67 Index, n=68	28.73±26.834	17.31±13.337	-	p=0.022*
Oncotype DX Recurrence Score®, n=74	26.52±9.774	14.63±6.174	-	p<0.001*
Magee Equation Recurrence Score, n=119	24.65±9.009	18.50±5.707	-	p<0.001*

Note. *=p<0.05; AO=Anastrozole Only; BDI-II=Beck Depression Inventory-II; BPI=Brief Pain Inventory; CA=Chemotherapy plus Anastrozole; ER=Estrogen Receptor; HER2= Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; LV=Lymphovascular; NART-R=National Adult Reading Test-Revised; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; POMS=Profile of Mood States; PR=Progesterone Receptor; SD=Standard Deviation; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors. [a]One-way ANOVAs utilized to compare means of continuous variables. [b]Pearson's Chi-square Tests of independence, Fisher's Exact Test, or Fisher's Exact Test computed using 2-sided Monte Carlo sampling based on 10,000 sampled tables used to examine associations between categorical variables. Only participants with complete confounder/covariate information are included in the participant characteristic statistics.

Table 16: SNPs Included in Genetic Regression Analyses (N=220)

Gene SNP	Wildtype/Variant Allele[a]	n for each SNP	Study MAF	HWE[b]
AURKA				
rs1047972	G/A	219	0.148	p=1.0000
rs16979877	A/G	207	0.080	p=0.3725
rs2273535	A/T	213	0.216	p=0.9788
rs6064389	G/T	219	0.445	p=0.2279
BAG1				
rs706118	T/G	214	0.248	p=0.1553
BCL2				
rs1564483	G/A	207	0.271	p=0.6855
rs17759659	A/G	219	0.425	p=0.8885
rs2279115	A/C	206	0.459	p=0.7047
rs4941195	C/A	218	0.429	p=0.4230
rs4987852	A/G	218	0.078	p=1.0000
rs4987853	A/G	217	0.189	p=0.5787
rs4987855	G/A	220	0.071	p=0.6073
rs956572	G/A	211	0.398	p=0.6542
rs9807663	T/A	218	0.108	p=0.7216
BIRC5				
rs1042489	T/C	217	0.359	p=0.5481
rs1508147	G/A	218	0.358	p=0.5738
rs17878467	C/T	217	0.111	p=1.0000
rs2239680	T/C	214	0.299	p=0.7100
rs3764383	A/G	213	0.305	p=0.9577
rs8073069	G/C	207	0.249	p=0.9445
rs8073903	T/C	212	0.366	p=0.6934
rs9904341	G/C	206	0.318	p=0.7061
CCNB1				
rs164390	G/T	214	0.371	p=0.5200
rs350099	T/C	216	0.396	p=0.5396
rs350104	T/C	219	0.459	p=0.5640
CD68				
rs8066665	G/A	220	0.457	p=0.1667
rs9901673	C/A	218	0.172	p=0.0915
CENPA				
rs3806517	A/G	215	0.340	p=0.8111
rs3806518	T/C	214	0.278	p=0.8532
CMC2				
rs1025065	C/A	209	0.361	p=0.8277
rs1981867	C/T	220	0.307	p=0.1739
rs9936489	T/G	215	0.319	p=0.7119
CTSL2				
rs16919034	A/G	213	0.169	p=0.3501
rs4361859	A/G	219	0.327	p=0.0078* p=0.0695[HC]
DIAPH3				
rs1337652	G/A	217	0.212	p=0.9194
rs4547237	A/G	220	0.307	p=0.8219
ERBB2				
rs1058808	G/C	217	0.373	p=0.2746

rs1136201	A/G	220	0.232	p=0.9465
rs1476278	A/G	220	0.357	p=0.9976
rs1810132	T/C	212	0.318	p=0.1537
rs2517955	T/C	220	0.373	p=0.4823
rs4252596	C/A	220	0.121	p=0.7482
rs903501	G/A	211	0.332	p=0.3885
rs9303274	C/T	219	0.356	p=0.9485
rs12976445 (*MIR125A*)	T/C	220	0.298	p=0.1467
ESR1				
rs10484919	C/T	204	0.088	p=0.6574
rs1062577	T/A	215	0.081	p=0.1495
rs11964281	C/T	214	0.075	p=1.0000
rs12173570	C/T	219	0.132	p=0.3857
rs12665044	C/T	213	0.132	p=0.7665
rs1514348	A/C	220	0.468	p=0.9519
rs1801132	C/G	220	0.232	p=0.4098
rs1884051	A/G	207	0.336	p=0.9170
rs2046210	C/T	213	0.357	p=0.3520
rs2071454	T/G	213	0.110	p=1.0000
rs2077647	G/A	217	0.484	p=0.0643
rs2228480	G/A	218	0.188	p=0.2292
rs2234693	C/T	206	0.481	p=0.0344* p=0.9104^{HC}
rs2347867	A/G	215	0.340	p=0.5008
rs2744677	A/C	215	0.249	p=0.6319
rs2813543	G/A	213	0.181	p=0.9848
rs2813544	A/G	215	0.235	p=0.0512
rs2941740	T/C	220	0.391	p=0.4585
rs3020314	T/C	208	0.358	p=0.4848
rs3778099	T/C	208	0.089	p=1.0000
rs3798577	T/C	220	0.475	p=0.2770
rs488133	C/T	209	0.285	p=0.5101
rs532010	T/C	212	0.434	p=0.7636
rs6557171	C/T	218	0.303	p=0.7528
rs77275268	C/T	217	0.083	p=0.1827
rs7761133	T/C	216	0.153	p=0.5840
rs7761846	T/C	204	0.054	p=0.4501
rs7766585	T/G	218	0.154	p=0.1375
rs7767143	A/G	214	0.243	p=0.0849
rs827421	C/T	212	0.467	p=0.1489
rs851967	G/A	216	0.308	p=0.6380
rs851971	G/A	216	0.313	p=0.5073
rs851982	T/C	217	0.362	p=0.3178
rs851998	C/T	219	0.313	p=0.6540
rs910416	T/C	220	0.491	p=0.9961
rs9322331	C/T	215	0.381	p=0.8339
rs9340799	A/G	213	0.397	p=0.4781
rs9383938	G/T	218	0.083	p=0.3721
rs9397435	A/G	220	0.073	p=0.6108
rs9397456	G/A	203	0.217	p=0.1522
rs985694	C/T	218	0.120	p=0.7731
rs1038304 (*CCDC170*)	G/A	218	0.456	p=0.1394
rs12662670 (*CCDC170*)	T/G	217	0.069	p=0.2740
rs3734805 (*CCDC170*)	A/C	213	0.075	p=0.2359

rs3757318 (*CCDC170*)	G/A	213	0.059	p=0.3629
rs6929137 (*CCDC170*)	G/A	216	0.319	p=0.3411
GRB7				
rs9910678	T/C	218	0.053	p=0.1079
GSTM1				
rs1065411	C/G	209	0.194	p=0.0884
rs412543	G/C	216	0.081	p=0.3714
rs35652124 (*NFE2L2*)	T/C	218	0.298	p=0.2739
rs6721961 (*NFE2L2*)	G/T	212	0.101	p=1.0000
MELK				
rs10973007	C/G	209	0.189	p=0.8337
rs2250340	C/T	220	0.075	p=1.0000
rs3780350	C/T	213	0.155	p=0.6424
MKI67				
rs10732438	A/G	211	0.367	p=0.1859
rs10764751	A/C	220	0.239	p=0.5706
MMP11				
rs131451	T/C	216	0.107	p=1.0000
MYBL2				
rs11556379	C/G	220	0.050	p=0.4243
rs2070235	A/G	220	0.093	p=1.0000
rs619289	C/T	216	0.197	p=0.7837
rs826943	T/C	213	0.146	p=0.7823
rs826944	C/T	219	0.142	p=1.0000
NDC80				
rs12408485	A/G	203	0.382	p=0.4731
rs2292274	T/C	207	0.268	p=0.5054
ORC6				
rs33994299	T/C	220	0.475	p=0.0051* $p=0.1405^{HC}$
PGR				
rs1042838	G/T	216	0.141	p=0.0466* $p=0.0160^{HC}$*
rs1042839	C/T	208	0.130	p=0.0103* $p=0.0027^{HC}$*
rs10895068	G/A	214	0.063	p=1.0000
rs11224561	C/T	214	0.119	p=0.7460
rs1893505	C/T	220	0.382	p=0.7593
rs1942836	T/C	217	0.201	p=0.7604
rs471767	A/G	216	0.313	p=0.3574
rs474320	T/A	197	0.147	p=0.0434* $p=0.0329^{HC}$*
rs4754732	T/C	220	0.334	p=0.1780
rs484389	T/C	212	0.217	p=0.6804
rs568157	A/G	219	0.493	p=0.3120
rs590688	C/G	215	0.463	p=0.1662
rs608995	A/T	218	0.220	p=0.5724
RACGAP1				
rs7303531	G/A	214	0.058	p=1.0000
RFC4				
rs1354091	A/C	214	0.238	p=0.9537
RRM2				
rs1138729	A/G	202	0.136	p=0.3821
rs4309551	C/T	218	0.452	p=0.8925

rs4668664	G/A	215	0.263	p=0.6850
SCUBE2				
rs1136966	T/G	213	0.211	p=0.8348
rs4910440	C/T	219	0.470	p=0.4879
rs6486125	A/G	207	0.266	p=0.1980

Note. *=p<0.05; HC=HWE p-value in healthy control participants only; HWE=Hardy-Weinberg Equilibrium; MAF=minor allele frequency; SNP=Single Nucleotide Polymorphism. [a]Wildtype and variant alleles based on study sample. [b]Chi-square Goodness-of-fit or Exact Test p-value.

Table 17: GRS and Cognitive Function Composite Regression Analysis Results

Cognitive Function Composite	Gene-SNP used in GRS Calculation	b_{GRS}, p-value[a]	Model R^2	R^2 Change for GRS
Attention (n=201)	*ERBB2*(*MIR125A*)-rs12976445 *ESR1*-rs2347867 *ESR1*-rs3020314 *ESR1*-rs6557171 *ESR1*-rs985694 *MYBL2*-rs2070235 *SCUBE2*-rs6486125	0.4665, p<0.001* 0.4800, p<0.001*	0.2593	0.066
Concentration (n=177)	*AURKA*-rs1047972 *BCL2*-rs9807663 *CCNB1*-rs164390 *CCNB1*-rs350099 *CENPA*-rs3806517 *DIAPH3*-rs4547237 *ESR1*-rs488133 *ESR1*-rs7767143 *ESR1*-rs910416 *ESR1*-rs9397456 *ESR1*(*CCDC170*)-rs12662670 *ESR1*(*CCDC170*)-rs3734805 *ESR1*(*CCDC170*)-rs3757318 *ESR1*(*CCDC170*)-rs6929137 *GRB7*-rs9910678 *MELK*-rs10973007 *PGR*-rs10895068	0.5098, p<0.001* 0.5358, p<0.001*	0.2495	0.189
Executive Function (n=137)	*BAG1*-rs706118 *BCL2*-rs1564483 *BCL2*-rs4987853 *CCNB1*-rs164390 *CCNB1*-rs350099 *CCNB1*-rs350104 *CTSL2*-rs4361859 *DIAPH3*-rs1337652 *DIAPH3*-rs4547237 *ESR1*-rs2234693 *ESR1*-rs488133 *ESR1*-rs7761846 *ESR1*-rs827421 *ESR1*(*CCDC170*)-rs3757318 *GSTM1*-rs412543 *MELK*-rs10973007 *MELK*-rs2250340 *MYBL2*-rs11556379 *PGR*-rs1042838 *PGR*-rs474320 *PGR*-rs484389 *PGR*-rs608995 *SCUBE2*-rs6486125	0.3526, p<0.001* 0.3589, p<0.001*	0.4296	0.204
Mental Flexibility (n=154)	*BCL2*-rs1564483 *BCL2*-rs4987853 *DIAPH3*-rs1337652 *ERBB2*(*MIR125A*)rs12976445	0.5040, p<0.001* 0.5383,	0.4712	0.224

	ESR1-rs2347867 *ESR1*-rs6557171 *ESR1*-rs985694 *GSTM1*(*NFE2L2*)-rs35652124 *MKI67*-rs10732438 *MYBL2*-rs11556379 *NDC80*-rs12408485 *NDC80*-rs2292274 *RFC4*-rs1354091 *RRM2*-rs1138729 *SCUBE2*-rs6486125	p,0.001*		
Psychomotor Speed (n=181)	*BCL2*-rs4941195 *BCL2*-rs956572 *CENPA*-rs3806518 *ESR1*-rs2347867 *ESR1*-rs488133 *ESR1*-rs9322331 *ESR1*-rs9340799 *MKI67*-rs10732438 *PGR*-rs568157	0.7265, p<0.001* 0.6674, p<0.001*	0.2527	0.093
Verbal Memory (n=146)	*AURKA*-rs16979877 *BCL2*-rs2279115 *BCL2*-rs4987852 *BIRC5*-rs3764383 *CCNB1*-rs164390 *CCNB1*-rs350099 *CCNB1*-rs350104 *CD68*-rs9901673 *CENPA*-rs3806518 *CTSL2*-rs16919034 *DIAPH3*-rs4547237 *ESR1*-rs10484919 *ESR1*-rs12665044 *ESR1*-rs2941740 *ESR1*-rs488133 *ESR1*-rs77275268 *ESR1*-rs7767143 *ESR1*-rs9383938 *ESR1*-rs9397435 *ESR1*(*CCDC170*)-rs3734805 *ESR1*(*CCDC170*)-rs3757318 *GSTM1*-rs412543 *MYBL2*-rs2070235 *MYBL2*-rs619289 *NDC80*-rs2292274 *ORC6*-rs33994299 *PGR*-rs484389 *PGR*-rs568157	0.3406, p<0.001* 0.3401, p<0.001*	0.5048	0.209
Visual Memory (n=165)	*BAG1*-rs706118 *BCL2*-rs1564483 *CCNB1*-rs350104 *DIAPH3*-rs1337652 *DIAPH3*-rs4547237 *ESR1*-rs2077647 *ESR1*-rs2813544 *ESR1*-rs488133	0.7477, p<0.001* 0.6078, p<0.001*	0.3167	0.148

	ESR1-rs7761846 *ESR1*-rs7767143 *ESR1*(*CCDC170*)-rs3757318 *GSTM1*-rs412543 *MYBL2*-rs2070235 *PGR*-rs11224561 *PGR*-rs1942836 *RRM2*-rs4309551			
Visual Working Memory (n=154)	*AURKA*-rs2273535 *BAG1*-rs706118 *BIRC5*-rs1508147 *BIRC5*-rs9904341 *CCNB1*-rs164390 *CCNB1*-rs350099 *CCNB1*-rs350104 *CD68*-rs9901673 *DIAPH3*-rs1337652 *DIAPH3*-rs4547237 *ESR1*-rs2941740 *ESR1*-rs488133 *ESR1*-rs7761846 *ESR1*-rs910416 *ESR1*-rs9397456 *GRB7*-rs9910678 *GSTM1*-rs412543 *MELK*-rs2250340 *MYBL2*-rs2070235 *MYBL2*-rs619289 *PGR*-rs11224561 *PGR*-rs608995	0.4198, p<0.001* 0.4131, p<0.001*	0.4700	0.241

Note. *=p<0.001; GRS=Genetic Risk/Protection Score. [a]Standard multiple linear regression coefficient and p-value listed first, robust multiple linear regression coefficient and p-value listed subsequently. Model R^2 and R^2 change reported from standard multiple linear regression models. Participants missing genetic data necessary for completion of a GRS calculation were not included in the GRS analysis. All regression models are adjusted for age, estimated verbal intelligence, levels of depressive symptoms, anxiety, fatigue, and pain, and prescribed treatment group.

Table 18: Individual SNP and Cognitive Function Robust Regression Results

Regression coefficient, p-value	Attention	Concentration	Executive Function	Mental Flexibility	Psychomotor Speed	Verbal Memory	Visual Memory	Visual Working Memory
***AURKA*rs1047972**	n=218	n=218	n=219	n=219	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	0.000, 0.998	-0.148, 0.241	0.080, 0.389	0.017, 0.867	0.149, 0.153	0.025, 0.806	0.021, 0.735	0.116, 0.263
Interaction Model								
SNP Main Effect	0.125, 0.579	-0.488, 0.016*	0.148, 0.353	0.161, 0.362	0.063, 0.716	-0.092, 0.601	0.152, 0.149	0.054, 0.762
SNP-by-CA Inter	-0.176, 0.617	0.569, 0.073	-0.073, 0.766	-0.407, 0.138	-0.075, 0.781	0.287, 0.292	-0.177, 0.275	0.119, 0.664
SNP-by-AO Inter	-0.205, 0.504	0.513. 0.063	-0.111, 0.609	-0.066, 0.784	0.298, 0.213	0.103, 0.669	-0.212, 0.140	0.067, 0.783
***AURKA*rs16979877**	n=206	n=206	n=207	n=206	n=207	n=207	n=207	n=207
Main Effects Model								
SNP Main Effect	-0.211, 0.238	0.125, 0.432	-0.147, 0.233	-0.094, 0.497	-0.199, 0.161	-0.138, 0.318	-0.157, 0.062	-0.135, 0.330
Interaction Model								
SNP Main Effect	-0.274, 0.391	0.180, 0.533	-0.234, 0.302	0.085, 0.735	0.035, 0.893	-0.540, 0.029*	-0.291, 0.056	-0.350, 0.166
SNP-by-CA Inter	0.295, 0.626	0.076, 0.879	-0.061, 0.875	-0.446, 0.305	-0.457, 0.305	0.577, 0.174	0.259, 0.322	0.496, 0.253
SNP-by-AO Inter	0.042, 0.916	-0.108, 0.763	0.176, 0.532	-0.184, 0.557	-0.354, 0.271	0.548, 0.074	0.152, 0.422	0.238, 0.447
***AURKA*rs2273535**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	0.040, 0.748	-0.063, 0.566	0.038, 0.663	-0.136, 0.160	-0.021, 0.834	0.186, 0.052	-0.064, 0.273	0.009, 0.926
Interaction Model								
SNP Main Effect	-0.078, 0.698	-0.027, 0.877	0.193, 0.164	-0.200, 0.198	-0.010, 0.949	0.297, 0.051	-0.082, 0.397	0.183, 0.241
SNP-by-CA Inter	0.198, 0.545	-0.104, 0.715	-0.146, 0.520	0.218, 0.387	0.282, 0.267	-0.031, 0.899	0.147, 0.354	-0.008, 0.975
SNP-by-AO Inter	0.186, 0.517	-0.019, 0.939	-0.306, 0.131	0.011, 0.961	-0.246, 0.271	-0.275, 0.208	-0.159, 0.254	-0.477, 0.034*
***AURKA*rs6064389**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	0.027, 0.839	0.067, 0.572	0.089, 0.341	0.054, 0.608	0.011, 0.915	-0.051, 0.623	-0.031, 0.610	0.178, 0.088
Interaction Model								
SNP Main Effect	-0.065, 0.769	0.119, 0.540	0.035, 0.819	-0.044, 0.797	-0.080, 0.637	-0.208, 0.220	0.050, 0.627	0.060, 0.724
SNP-by-CA Inter	0.557, 0.189	-0.292. 0.414	0.202, 0.474	0.176, 0.578	0.562, 0.072	0.665, 0.831	-0.151, 0.426	-0.022, 0.945
SNP-by-AO Inter	0.080, 0.788	-0.005, 0.986	0.046, 0.823	0.140. 0.548	-0.004, 0.986	0.303, 0.186	-0.106, 0.448	0.267, 0.248
***BAG1*rs706118**	n=213	n=213	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								

SNP Main Effect	-0.101, 0.408	-0.109, 0.314	-0.123, 0.145	0.045, 0.624	-0.016, 0.870	-0.154, 0.103	-0.107, 0.067	-0.310, 0.001*
Interaction Model								
SNP Main Effect	0.061, 0.763	0.069, 0.706	-0.312, 0.027*	0.125, 0.441	0.006, 0.971	-0.127, 0.421	-0.216, 0.025*	-0.450, 0.004*
SNP-by-CA Inter	-0.549, 0.092	-0.270. 0.355	0.263, 0.239	-0.213, 0.406	-0.324, 0.209	-0.109, 0.664	0.112, 0.465	0.099, 0.684
SNP-by-AO Inter	-0.064, 0.825	-0.288, 0.268	0.321, 0.108	-0.064, 0.781	0.174, 0.450	0.006, 0.980	0.219, 0.110	0.299, 0.170
***BCL2*rs1564483**	n=206	n=207	n=207	n=206	n=207	n=207	n=207	n=207
Main Effects Model								
SNP Main Effect	0.246, 0.052	0.001, 0.994	0.191, 0.030*	0.049, 0.621	0.082, 0.410	0.067, 0.491	0.118, 0.221[b]	0.085, 0.389
Interaction Model								
SNP Main Effect	0.312, 0.141	-0.224, 0.247	0.397, 0.007*	0.396, 0.017*	0.167, 0.321	0.246, 0.134	0.226, 0.031*	0.244, 0.147
SNP-by-CA Inter	-0.256, 0.424	0.185, 0.526	-0.244, 0.267	-0.607, 0.015*	-0.326, 0.200	-0.317, 0.200	-0.386, 0.015*	-0.411, 0.105
SNP-by-AO Inter	0.023, 0.937	0.461, 0.085	-0.352, 0.081	-0.406, 0.074	0.021, 0.928	-0.222, 0.328	-0.107, 0.457	-0.068, 0.770
***BCL2*rs17759659**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	0.132, 0.311	0.093, 0.415	0.039, 0.671	0.080, 0.434	-0.003, 0.973	0.043, 0.668	-0.004, 0.951	-0.043, 0.674
Interaction Model								
SNP Main Effect	0.320, 0.111	0.098, 0.584	0.069, 0.625	0.047, 0.769	0.104, 0.512	0.223, 0.152	-0.036, 0.707	-0.176, 0.268
SNP-by-CA Inter	-0.526, 0.108	0.114, 0.696	-0.131, 0.572	0.164, 0.527	0.033, 0.899	-0.456, 0.072	0.106, 0.502	0.193, 0.457
SNP-by-AO Inter	-0.173, 0.548	-0.098, 0.703	0.003, 0.988	-0.029, 0.901	-0.334, 0.146	-0.174, 0.437	0.014, 0.920	0.235, 0.305
***BCL2*rs2279115**	n=205	n=205	n=206	n=206	n=206	n=206	n=206	n=206
Main Effects Model								
SNP Main Effect	0.019, 0.888	-0.140, 0.251	0.033, 0.722	-0.118, 0.266	-0.152, 0.147	-0.106, 0.320	-0.083, 0.197	-0.141, 0.191
Interaction Model								
SNP Main Effect	-0.033, 0.882	-0.169, 0.412	0.099, 0.531	-0.204, 0.252	-0.083, 0.640	-0.380, 0.034*	0.011, 0.921	-0.097, 0.594
SNP-by-CA Inter	-0.039, 0.910	-0.008, 0.980	-0.063, 0.797	0.070, 0.797	-0.165, 0.544	0.348, 0.207	-0.157, 0.357	-0.047, 0.867
SNP-by-AO Inter	0.188, 0.547	0.084, 0.771	-0.127, 0.568	0.180, 0.471	-0.073, 0.769	0.451, 0.073	-0.122, 0.434	-0.078, 0.761
***BCL2*rs4941195**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.049, 0.700	-0.095, 0.398	0.107, 0.222	-0.058, 0.560	-0.227, 0.021*	-0.003, 0.979	0.010, 0.864	0.013, 0.897
Interaction Model								
SNP Main Effect	-0.027, 0.895	-0.048, 0.789	0.214, 0.129	-0.110, 0.493	-0.185, 0.239	-0.213, 0.174	0.132, 0.169	0.300, 0.057
SNP-by-CA Inter	0.024, 0.941	-0.021, 0.942	-0.082, 0.719	0.096, 0.711	0.000, 1.000	0.158, 0.534	-0.221, 0.157	-0.471, 0.066
SNP-by-AO Inter	0.189, 0.514	-0.110, 0.668	-0.225, 0.264	0.074, 0.749	-0.123, 0.586	0.434, 0.055	-0.174, 0.208	-0.407, 0.073

***BCL2*rs4987852**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.273, 0.096	-0.154, 0.297	-0.072, 0.543	-0.068, 0.603	-0.058, 0.656	0.190, 0.140	-0.018, 0.818	-0.008, 0.953
Interaction Model								
SNP Main Effect	-0.511, 0.051	0.046, 0.849	-0.070, 0.709	-0.097, 0.652	0.087, 0.681	0.464, 0.024*	-0.005, 0.965	-0.121, 0.566
SNP-by-CA Inter	0.458, 0.336	-0.528, 0.226	0.282, 0.409	0.079, 0.837	-0.219, 0.570	-0.046, 0.902	0.214, 0.350	0.382, 0.320
SNP-by-AO Inter	0.383, 0.284	-0.178, 0.586	-0.101, 0.693	0.030, 0.917	-0.310, 0.284	-0.542, 0.054	-0.120, 0.486	0.076, 0.792
***BCL2*rs4987853**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	-0.033, 0.793	-0.046, 0.681	-0.035, 0.687	-0.053, 0.589	0.014, 0.887	-0.025, 0.796	-0.062, 0.314	-0.098, 0.322
Interaction Model								
SNP Main Effect	-0.182, 0.386	-0.045, 0.811	-0.282, 0.054	-0.313, 0.058	-0.134, 0.418	-0.253, 0.121	-0.078, 0.448	-0.145, 0.384
SNP-by-CA Inter	0.239, 0.492	0.179, 0.560	0.516, 0.031*	0.211, 0.430	0.205, 0.451	0.378, 0.157	0.141, 0.403	0.149, 0.583
SNP-by-AO Inter	0.234, 0.423	-0.113, 0.662	0.304, 0.134	0.497, 0.029*	0.255, 0.269	0.329, 0.147	-0.055, 0.701	0.022, 0.924
***BCL2*rs4987855**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.064, 0.708	-0.020, 0.898	-0.055, 0.651	-0.101, 0.453	0.004, 0.975	0.029, 0.826	-0.086, 0.296	-0.111, 0.408
Interaction Model								
SNP Main Effect	-0.033, 0.907	0.093, 0.712	-0.306, 0.118	-0.293, 0.182	-0.115, 0.604	-0.156, 0.474	-0.079, 0.558	-0.304, 0.168
SNP-by-CA Inter	0.008, 0.987	-0.334, 0.487	0.255, 0.496	0.283, 0.499	0.037, 0.931	0.363, 0.383	-0.061, 0.811	0.228, 0.589
SNP-by-AO Inter	0.209, 0.581	-0.149, 0.659	0.433, 0.101	0.298, 0.313	0.269, 0.367	0.254, 0.387	0.011, 0.949	0.295, 0.321
***BCL2*rs956572**	n=210	n=211	n=211	n=210	n=211	n=211	n=211	n=211
Main Effects Model								
SNP Main Effect	0.081, 0.532	0.039, 0.725	-0.028, 0.757	0.049, 0.626	0.170, 0.091	-0.169, 0.092	-0.015, 0.809	0.013, 0.897
Interaction Model								
SNP Main Effect	0.222, 0.311	-0.089, 0.638	0.093, 0.545	0.278, 0.098	0.392, 0.020*	-0.191, 0.257	0.085, 0.426	-0.009, 0.960
SNP-by-CA Inter	-0.219, 0.524	0.122, 0.680	-0.181, 0.451	-0.171, 0.512	-0.283, 0.280	-0.097, 0.712	-0.046, 0.782	-0.007, 0.979
SNP-by-AO Inter	-0.223, 0.454	0.241, 0.351	-0.187, 0.374	-0.441, 0.054	-0.411, 0.073	0.108, 0.637	-0.250, 0.084	0.064, 0.788
***BCL2*rs9807663**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.012, 0.938	0.103, 0.439	0.012, 0.907	-0.011, 0.927	0.033, 0.779	-0.065, 0.572	0.044, 0.528	-0.043, 0.716
Interaction Model								
SNP Main Effect	0.028, 0.899	0.315, 0.107	0.107, 0.485	-0.027, 0.874	0.210, 0.223	-0.235, 0.165	0.042, 0.679	-0.001, 0.995
SNP-by-CA Inter	-0.516, 0.196	-0.178, 0.617	-0.349, 0.198	-0.367, 0.222	-0.431, 0.156	0.189, 0.525	-0.230, 0.194	-0.213, 0.486

SNP-by-AO Inter	0.191, 0.578	-0.681, 0.028*	-0.039, 0.873	0.306, 0.256	-0.242. 0.372	0.349, 0.191	0.173, 0.276	0.059, 0.828
***BIRC5*rs1042489**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	-0.004, 0.977	0.080, 0.468	-0.015, 0.865	0.145, 0.129	-0.030, 0.758	-0.057, 0.552	-0.078, 0.174	-0.049, 0.609
Interaction Model								
SNP Main Effect	0.022, 0.911	0.176, 0.321	-0.094, 0.498	0.235, 0.127	-0.181, 0.247	-0.199, 0.194	-0.161, 0.082	-0.234, 0.125
SNP-by-CA Inter	-0.052, 0.869	-0.113, 0.690	0.102, 0.644	-0.199, 0.413	0.117, 0.635	0.178, 0.462	0.190, 0.195	0.451, 0.062
SNP-by-AO Inter	-0.035, 0.905	-0.199, 0.446	0.145, 0.479	-0.097, 0.667	0.356, 0.122	0.252, 0.263	0.091, 0.502	0.140, 0.533
***BIRC5*rs1508147**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.039, 0.750	0.067, 0.540	0.002, 0.983	0.138, 0.152	-0.025, 0.794	-0.032, 0.735	-0.061, 0.289	-0.052, 0.587
Interaction Model								
SNP Main Effect	0.033, 0.868	0.155, 0.383	-0.092, 0.507	0.226, 0.148	-0.187, 0.226	-0.190, 0.213	-0.155, 0.092	-0.249, 0.098
SNP-by-CA Inter	-0.013, 0.967	-0.089, 0.753	0.063, 0.773	-0.191, 0.436	0.146, 0.548	0.208, 0.385	0.206, 0.156	0.472, 0.048*
SNP-by-AO Inter	0.026, 0.929	-0.195, 0.459	0.211, 0.305	-0.090, 0.695	0.362, 0.115	0.271, 0.230	0.110, 0.421	0.132, 0.553
***BIRC5*rs17878467**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	-0.193, 0.184	0.246, 0.059	-0.121, 0.236	0.125, 0.278	-0.037, 0.748	-0.091, 0.416	-0.022, 0.753	-0.003, 0.979
Interaction Model								
SNP Main Effect	-0.413, 0.084	0.149, 0.489	-0.099, 0.562	0.063, 0.744	-0.192, 0.311	-0.008, 0.965	-0.106, 0.359	0.042, 0.826
SNP-by-CA Inter	0.351, 0.350	0.214, 0.520	-0.161, 0.537	0.001, 0.996	0.299, 0.304	-0.118, 0.682	0.140, 0.429	-0.008, 0.978
SNP-by-AO Inter	0.366, 0.285	0.091, 0.769	0.061, 0.801	0.176, 0.522	0.246, 0.364	-0.132, 0.623	0.148, 0.370	-0.125, 0.648
***BIRC5*rs2239680**	n=213	n=214	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	0.110, 0.385	-0.197, 0.073	0.042, 0.635	-0.086, 0.379	-0.056, 0.576	0.120, 0.219	0.068, 0.242	0.128, 0.197
Interaction Model								
SNP Main Effect	0.220, 0.275	-0.156, 0.367	-0.020, 0.889	-0.119, 0.448	-0.074, 0.637	0.101, 0.515	0.028, 0.766	0.001, 0.996
SNP-by-CA Inter	-0.199, 0.536	0.213, 0.437	0.100, 0.656	-0.022, 0.930	0.406, 0.103	0.027, 0.914	0.048, 0.749	0.270, 0.283
SNP-by-AO Inter	-0.162, 0.572	-0.269, 0.275	0.095, 0.637	0.098, 0.661	-0.235, 0.294	0.031, 0.890	0.083, 0.535	0.142, 0.530
***BIRC5*rs3764383**	n=212	n=212	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	0.149, 0.236	-0.165, 0.143	0.114, 0.199	-0.113, 0.250	-0.044, 0.662	0.192, 0.047*	0.107, 0.057	0.148, 0.140
Interaction Model								

SNP Main Effect	0.238, 0.235	-0.090, 0.614	-0.012, 0.931	-0.279, 0.078	-0.117, 0.462	0.183, 0.238	0.083, 0.369	-0.025, 0.875
SNP-by-CA Inter	-0.328, 0.293	0.094, 0.735	0.212, 0.338	0.122, 0.617	0.391, 0.112	0.030, 0.899	0.039, 0.786	0.337, 0.174
SNP-by-AO Inter	-0.008, 0.978	-0.268, 0.285	0.185, 0.356	0.346, 0.118	-0.089, 0.692	0.003, 0.991	0.062, 0.634	0.203, 0.366
***BIRC5*rs8073069**	n=206	n=206	n=207	n=206	n=207	n=207	n=207	n=207
Main Effects Model								
SNP Main Effect	0.094, 0.461	-0.091, 0.421	0.067, 0.444	0.042, 0.677	-0.025, 0.803	-0.028, 0.775	-0.099, 0.118	-0.134, 0.168
Interaction Model								
SNP Main Effect	0.338, 0.104	0.051, 0.780	0.056, 0.694	0.176, 0.287	-0.061, 0.707	-0.266, 0.097	-0.071, 0.488	-0.244, 0.119
SNP-by-CA Inter	-0.280, 0.416	-0.331, 0.278	-0.037, 0.874	-0.243, 0.371	-0.179, 0.508	0.266, 0.315	-0.022, 0.898	0.263, 0.308
SNP-by-AO Inter	-0.454, 0.131	-0.163, 0.538	0.044, 0.829	-0.184, 0.437	0.220, 0.353	0.442, 0.057	-0.061, 0.683	0.061, 0.786
***BIRC5*rs8073903**	n=211	n=212	n=212	n=211	n=212	n=212	n=212	n=212
Main Effects Model								
SNP Main Effect	0.047, 0.710	0.059, 0.586	0.025, 0.773	0.163, 0.092	0.005, 0.961	-0.030, 0.754	-0.065, 0.257	-0.025, 0.800
Interaction Model								
SNP Main Effect	0.048, 0.811	0.133, 0.452	-0.042, 0.766	0.291, 0.061	-0.145, 0.358	-0.215, 0.165	-0.133, 0.162	-0.163, 0.296
SNP-by-CA Inter	-0.068, 0.834	-0.109, 0.699	0.082, 0.718	-0.261, 0.292	0.135, 0.593	0.266, 0.285	0.201, 0.186	0.400, 0.111
SNP-by-AO Inter	0.046, 0.878	-0.131, 0.616	0.126, 0.549	-0.155, 0.500	0.349, 0.137	0.306, 0.183	0.042, 0.767	0.027, 0.908
***BIRC5*rs9904341**	n=205	n=205	n=206	n=205	n=206	n=206	n=206	n=206
Main Effects Model								
SNP Main Effect	-0.169, 0.172	0.064, 0.573	-0.021, 0.810	-0.060, 0.543	0.030, 0.762	-0.091, 0.350	-0.060, 0.320	-0.052, 0.592
Interaction Model								
SNP Main Effect	-0.173, 0.378	0.051, 0.773	0.171, 0.210	-0.045, 0.774	0.165, 0.292	0.078, 0.610	0.007, 0.943	0.139, 0.354
SNP-by-CA Inter	0.037, 0.910	-0.231, 0.436	-0.362, 0.109	0.036, 0.890	-0.381, 0.143	-0.279, 0.272	-0.091, 0.549	-0.510, 0.041*
SNP-by-AO Inter	-0.010, 0.971	0.190, 0.458	-0.283, 0.148	-0.063, 0.780	-0.145, 0.519	-0.258, 0.243	-0.129, 0.324	-0.150, 0.488
***CCNB1*rs164390**	n=213	n=214	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	0.010, 0.939	-0.230, 0.035*	-0.048, 0.582	-0.060, 0.538	-0.002, 0.986	0.014, 0.883	-0.010, 0.859	0.237, 0.016*
Interaction Model								
SNP Main Effect	0.196, 0.323	-0.167, 0.334	0.280, 0.039*	-0.116, 0.457	0.016, 0.921	0.299, 0.052	0.099, 0.298	0.581, <0.001*
SNP-by-CA Inter	-0.046, 0.884	-0.153, 0.573	-0.478, 0.026*	0.281, 0.252	0.054, 0.831	-0.355, 0.143	-0.147, 0.327	-0.549, 0.025*
SNP-by-AO Inter	-0.498, 0.085	-0.054, 0.829	-0.551, 0.006*	-0.067, 0.767	-0.097, 0.678	-0.518, 0.021*	-0.196, 0.156	-0.540, 0.017*
***CCNB1*rs350099**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								

SNP Main Effect	-0.013, 0.918	-0.246, 0.026*	-0.044, 0.618	-0.092, 0.343	0.030, 0.759	-0.027, 0.778	-0.050, 0.398	0.185, 0.059
Interaction Model								
SNP Main Effect	0.194, 0.332	-0.155, 0.377	0.236, 0.087	-0.125, 0.427	0.044, 0.782	0.350, 0.024*	0.069, 0.475	0.562, <0.001*
SNP-by-CA Inter	-0.132, 0.676	-0.162, 0.556	-0.345, 0.110	0.269, 0.273	0.001, 0.996	-0.445, 0.065	-0.118, 0.432	-0.553, 0.024*
SNP-by-AO Inter	-0.492, 0.089	-0.131, 0.606	-0.507, 0.011*	-0.121, 0.592	-0.036, 0.875	-0.685, 0.002*	-0.244, 0.079	-0.608, 0.008*
***CCNB1*rs350104**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	0.056, 0.677	0.053, 0.657	0.028, 0.763	0.144, 0.168	-0.102, 0.328	-0.069, 0.505	0.013, 0.843	-0.015, 0.889
Interaction Model								
SNP Main Effect	-0.178, 0.448	-0.023, 0.915	-0.225, 0.177	0.104, 0.575	-0.154, 0.401	-0.441, 0.016*	-0.153, 0.176	-0.475, 0.009*
SNP-by-CA Inter	0.174, 0.620	0.122, 0.702	0.128, 0.606	-0.001, 0.998	0.295, 0.282	0.564, 0.039*	0.193, 0.251	0.631, 0.020*
SNP-by-AO Inter	0.446, 0.150	0.106, 0.706	0.470, 0.032*	0.095, 0.698	-0.063, 0.794	0.516, 0.032*	0.297, 0.045*	0.699, 0.004*
***CD68*rs8066665**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.042, 0.741	-0.130, 0.250	0.021, 0.815	0.040, 0.688	-0.048, 0.630	-0.056, 0.574	0.012, 0.835	-0.044, 0.660
Interaction Model								
SNP Main Effect	-0.012, 0.959	-0.289, 0.142	0.120, 0.438	-0.027, 0.878	-0.142, 0.414	-0.252, 0.147	-0.084, 0.433	-0.220, 0.205
SNP-by-CA Inter	-0.159, 0.633	0.168, 0.570	-0.336, 0.149	-0.163, 0.534	0.023, 0.928	0.364, 0.163	0.143, 0.374	0.275, 0.292
SNP-by-AO Inter	0.235, 0.428	0.295, 0.263	-0.032, 0.876	0.273, 0.248	0.237, 0.311	0.250, 0.284	0.153, 0.289	0.247, 0.289
***CD68*rs9901673**	n=217	n=217	n=218	n=218	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.069, 0.595	0.005, 0.967	0.070, 0.453	0.013, 0.901	0.036, 0.733	0.098, 0.338	0.064, 0.306	0.100, 0.337
Interaction Model								
SNP Main Effect	0.176, 0.427	-0.020, 0.920	0.114, 0.458	0.078, 0.656	0.104, 0.559	0.446, 0.010*	0.128, 0.242	0.412, 0.019*
SNP-by-CA Inter	-0.202, 0.537	0.035, 0.905	0.315, 0.163	-0.273, 0.287	-0.169, 0.519	-0.444, 0.078	-0.191, 0.234	-0.439, 0.087
SNP-by-AO Inter	-0.141, 0.646	0.052, 0.853	-0.331, 0.121	0.038, 0.877	-0.060, 0.809	-0.555, 0.020*	-0.011, 0.940	-0.498, 0.040*
***CENPA*rs3806517**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	0.028, 0.817	-0.127, 0.252	-0.021, 0.801	-0.002, 0.980	0.009, 0.922	-0.020, 0.831	-0.004, 0.951	0.070, 0.457
Interaction Model								
SNP Main Effect	-0.229, 0.247	-0.363, 0.039*	0.107, 0.443	0.008, 0.958	-0.118, 0.449	-0.196, 0.205	-0.013, 0.890	0.074, 0.622
SNP-by-CA Inter	0.237, 0.471	0.056, 0.849	-0.143, 0.539	0.039, 0.881	0.291, 0.264	0.353, 0.171	-0.024, 0.879	0.283, 0.259
SNP-by-AO Inter	0.538, 0.052	0.615, 0.012*	-0.231, 0.238	-0.050, 0.820	0.171, 0.432	0.226, 0.296	0.042, 0.750	-0.193, 0.357

***CENPA*rs3806518**	n=213	n=213	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	0.011, 0.927	0.017, 0.877	-0.005, 0.956	-0.044, 0.639	0.050, 0.601	-0.019, 0.838	-0.023, 0.684	0.025, 0.792
Interaction Model								
SNP Main Effect	0.136, 0.493	-0.103, 0.561	-0.101, 0.457	-0.195, 0.211	-0.148, 0.335	0.282, 0.063	0.046, 0.630	0.139, 0.369
SNP-by-CA Inter	-0.395, 0.212	0.312, 0.270	0.011, 0.960	0.368, 0.135	0.079, 0.745	-0.501, 0.037*	-0.086, 0.572	-0.340, 0.167
SNP-by-AO Inter	-0.058, 0.834	0.078, 0.753	0.226, 0.230	0.130, 0.548	0.507, 0.018*	-0.423, 0.045*	-0.140, 0.294	-0.047, 0.828
***CMC2*rs1025065**	n=208	n=208	n=209	n=208	n=209	n=209	n=209	n=209
Main Effects Model								
SNP Main Effect	-0.128, 0.298	0.057, 0.603	0.002, 0.983	0.019, 0.845	0.178, 0.175	-0.109, 0.257	0.046, 0.421	-0.022, 0.818
Interaction Model								
SNP Main Effect	-0.217, 0.291	0.135, 0.463	0.038, 0.793	0.019, 0.907	0.084, 0.607	-0.218, 0.172	0.032, 0.741	-0.088, 0.580
SNP-by-CA Inter	-0.141, 0.660	-0.048, 0.868	-0.275, 0.216	-0.138, 0.581	0.144, 0.566	0.260, 0.294	0.026, 0.865	0.190, 0.444
SNP-by-AO Inter	0.363, 0.216	-0.195, 0.459	0.088, 0.667	0.107, 0.641	-0.190, 0.412	0.099, 0.663	0.015, 0.913	0.029, 0.898
***CMC2*rs1981867**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.136, 0.253	-0.197, 0.058	0.022, 0.792	0.059, 0.526	-0.091, 0.326	0.060, 0.512	0.061, 0.279	0.094, 0.310
Interaction Model								
SNP Main Effect	-0.221, 0.262	0.018, 0.914	0.069, 0.617	0.147, 0.342	-0.031, 0.839	0.172, 0.258	0.020, 0.832	0.009, 0.951
SNP-by-CA Inter	0.153, 0.621	-0.300, 0.264	0.045, 0.835	-0.227, 0.350	0.035, 0.884	-0.182, 0.447	0.073, 0.617	-0.024, 0.920
SNP-by-AO Inter	0.121, 0.662	-0.363, 0.127	-0.137, 0.478	-0.069, 0.751	-0.201, 0.354	-0.160, 0.453	0.061, 0.640	0.239, 0.265
***CMC2*rs9936489**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	0.151, 0.213	-0.113, 0.302	-0.026, 0.757	-0.048, 0.618	-0.037, 0.705	-0.071, 0.451	0.030, 0.600	-0.057, 0.544
Interaction Model								
SNP Main Effect	0.264, 0.183	-0.186, 0.300	-0.147, 0.293	0.051, 0.743	0.065, 0.676	-0.135, 0.377	-0.008, 0.932	-0.061, 0.693
SNP-by-CA Inter	-0.131, 0.681	0.082, 0.777	0.281, 0.209	-0.115, 0.646	-0.365, 0.146	-0.053, 0.828	0.066, 0.662	-0.149, 0.545
SNP-by-AO Inter	-0.220, 0.437	0.147, 0.567	0.141, 0.481	-0.185, 0.408	-0.013, 0.953	0.197, 0.367	0.054, 0.689	0.127, 0.563
***CTSL2*rs16919034**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	0.112, 0.397	-0.058, 0.611	-0.027, 0.768	-0.090, 0.378	0.002, 0.987	-0.120, 0.242	-0.007, 0.909	0.112, 0.286
Interaction Model								
SNP Main Effect	0.007, 0.974	0.226, 0.236	-0.243, 0.110	-0.094, 0.582	-0.023, 0.894	-0.429, 0.011*	-0.043, 0.678	0.046, 0.791
SNP-by-CA Inter	0.339, 0.320	-0.309, 0.295	0.216, 0.357	0.038, 0.886	-0.084, 0.756	0.134, 0.603	0.000, 0.998	-0.147, 0.582

SNP-by-AO Inter	0.046, 0.882	-0.529, 0.054	0.408, 0.061	-0.017, 0.944	0.149, 0.552	0.696, 0.004*	0.109, 0.462	0.290, 0.243
***CTSL2*rs4361859**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	-0.029, 0.807	-0.029, 0.787	0.025, 0.766	0.068, 0.460	-0.002, 0.987	0.007, 0.942	-0.027, 0.622	-0.045, 0.627
Interaction Model								
SNP Main Effect	0.016, 0.934	0.066, 0.709	0.272, 0.045*	0.044, 0.771	0.077, 0.618	0.162, 0.287	0.094, 0.315	-0.072, 0.639
SNP-by-CA Inter	-0.505, 0.099	-0.176, 0.531	-0.433, 0.044*	-0.240, 0.318	-0.234, 0.343	-0.207, 0.392	-0.173, 0.242	0.069, 0.779
SNP-by-AO Inter	0.274, 0.307	-0.115, 0.640	-0.341, 0.072	0.252, 0.233	-0.038, 0.862	-0.255, 0.232	-0.207, 0.114	0.015, 0.945
***DIAPH3*rs1337652**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	-0.069, 0.582	0.130, 0.247	-0.163, 0.063	0.018, 0.857	-0.141, 0.157	-0.048, 0.624	-0.126, 0.032*	-0.090, 0.365
Interaction Model								
SNP Main Effect	-0.049, 0.807	0.006, 0.973	-0.295, 0.036*	0.362, 0.021*	-0.161, 0.313	0.037, 0.815	-0.187, 0.048*	0.095, 0.550
SNP-by-CA Inter	0.456, 0.174	0.093, 0.759	0.405, 0.084	-0.361, 0.166	0.333, 0.212	0.040, 0.880	0.305, 0.055	0.002, 0.995
SNP-by-AO Inter	-0.312, 0.255	0.268, 0.281	0.129, 0.503	-0.620, 0.004*	-0.135, 0.539	-0.213, 0.329	-0.054, 0.676	-0.446, 0.042*
***DIAPH3*rs4547237**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.147, 0.212	-0.035, 0.738	0.185, 0.024*	0.023, 0.802	0.157, 0.088	0.198, 0.030*	0.084, 0.127	0.195, 0.034*
Interaction Model								
SNP Main Effect	0.081, 0.679	-0.273, 0.118	0.410, 0.002*	-0.130, 0.406	0.095, 0.537	0.355, 0.018*	0.194, 0.039*	0.229, 0.134
SNP-by-CA Inter	0.102, 0.742	0.053, 0.847	-0.280, 0.184	0.208, 0.394	0.090, 0.709	-0.090, 0.702	-0.109, 0.456	-0.129, 0.593
SNP-by-AO Inter	0.105, 0.704	0.610, 0.014*	-0.389, 0.040*	0.234, 0.285	0.114, 0.597	-0.337, 0.111	-0.225, 0.088	0.002, 0.991
***ERBB2*rs1058808**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	0.059, 0.627	0.030, 0.784	-0.023, 0.788	0.041, 0.668	0.044, 0.642	0.035, 0.711	0.001, 0.981	-0.061, 0.525
Interaction Model								
SNP Main Effect	-0.133, 0.507	0.079, 0.658	0.030, 0.829	0.050, 0.748	0.090, 0.573	0.169, 0.281	-0.064, 0.506	0.101, 0.523
SNP-by-CA Inter	0.360, 0.259	-0.105, 0.713	-0.137, 0.538	0.266, 0.280	0.192, 0.446	-0.081, 0.745	0.109, 0.470	-0.206, 0.409
SNP-by-AO Inter	0.266, 0.340	-0.058, 0.815	-0.045, 0.819	-0.195, 0.366	-0.261, 0.241	-0.275, 0.208	-0.560, 0.268	-0.270, 0.220
***ERBB2*rs1136201**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.016, 0.896	-0.034, 0.756	0.076, 0.371	-0.010, 0.920	-0.035, 0.716	-0.074, 0.435	-0.064, 0.265	-0.181, 0.056
Interaction Model								

SNP Main Effect	0.252, 0.205	0.159, 0.372	0.027, 0.847	-0.086, 0.588	-0.094, 0.549	-0.246, 0.111	-0.119, 0.203	-0.205, 0.187
SNP-by-CA Inter	-0.260, 0.411	-0.169, 0.553	0.239, 0.283	0.036, 0.885	0.108, 0.666	0.213, 0.384	0.119, 0.423	0.223, 0.367
SNP-by-AO Inter	-0.451, 0.116	-0.383, 0.135	-0.015, 0.940	0.163, 0.471	0.091, 0.687	0.296, 0.181	0.067, 0.615	-0.085, 0.703
***ERBB2*rs1476278**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.033, 0.783	-0.028, 0.797	-0.041, 0.629	-0.019, 0.838	0.025, 0.792	0.040, 0.665	0.016, 0.775	-0.023, 0.811
Interaction Model								
SNP Main Effect	-0.141, 0.485	0.090, 0.620	0.008, 0.952	0.042, 0.791	0.136, 0.385	0.130, 0.405	-0.064, 0.508	0.114, 0.471
SNP-by-CA Inter	0.334, 0.295	-0.126, 0.659	-0.256, 0.249	0.192, 0.437	0.032, 0.898	-0.027, 0.912	0.114, 0.456	-0.130, 0.597
SNP-by-AO Inter	0.228, 0.413	-0.227, 0.365	0.028, 0.885	-0.272, 0.211	-0.333, 0.126	-0.199, 0.359	0.138, 0.305	-0.255, 0.242
***ERBB2*rs1810132**	n=211	n=212	n=212	n=211	n=212	n=212	n=212	n=212
Main Effects Model								
SNP Main Effect	0.110, 0.371	0.097, 0.367	-0.001, 0.993	0.040, 0.670	0.124, 0.194	0.073, 0.442	0.015, 0.794	-0.013, 0.890
Interaction Model								
SNP Main Effect	-0.025, 0.901	0.191, 0.274	0.080, 0.562	0.164, 0.275	0.182, 0.242	0.114, 0.457	-0.048, 0.611	0.072, 0.648
SNP-by-CA Inter	0.319, 0.316	-0.070, 0.802	-0.136, 0.535	0.094, 0.694	0.128, 0.606	0.168, 0.492	0.177, 0.237	-0.071, 0.777
SNP-by-AO Inter	0.130, 0.644	-0.215, 0.386	-0.110, 0.575	-0.398, 0.063	-0.267, 0.229	-0.223, 0.307	0.035, 0.795	-0.183, 0.414
***ERBB2*rs2517955**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.022, 0.858	0.020, 0.850	-0.074, 0.383	0.011, 0.910	0.026, 0.786	0.018, 0.845	-0.011, 0.841	-0.047, 0.618
Interaction Model								
SNP Main Effect	-0.140, 0.486	0.086, 0.635	0.007, 0.960	0.040, 0.799	0.135, 0.390	0.130, 0.406	-0.058, 0.543	0.110, 0.483
SNP-by-CA Inter	0.347, 0.279	-0.137, 0.632	-0.273, 0.220	0.264, 0.285	0.024, 0.923	-0.068, 0.784	0.055, 0.715	-0.205, 0.408
SNP-by-AO Inter	0.191, 0.493	-0.075, 0.765	-0.042, 0.829	-0.241, 0.265	-0.320, 0.142	-0.226, 0.299	0.085, 0.519	-0.258, 0.236
***ERBB2*rs4252596**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.134, 0.343	-0.033, 0.791	0.033, 0.742	-0.005, 0.966	0.065, 0.553	0.078, 0.477	0.043, 0.515	0.039, 0.726
Interaction Model								
SNP Main Effect	-0.124, 0.564	-0.037, 0.850	0.056, 0.710	0.018, 0.919	-0.051, 0.759	0.000. 0.999	0.085, 0.403	0.082, 0.631
SNP-by-CA Inter	-0.169, 0.673	0.197, 0.568	0.059, 0.828	-0.430, 0.158	-0.085, 0.776	0.071, 0.812	-0.047, 0.797	-0.120, 0.692
SNP-by-AO Inter	0.076, 0.811	-0.110, 0.703	-0.091, 0.683	0.164, 0.511	0.376, 0.126	0.154, 0.531	-0.105, 0.482	-0.040, 0.872
***ERBB2*rs903501**	n=210	n=211	n=211	n=211	n=211	n=211	n=211	n=211
Main Effects Model								

SNP Main Effect	0.051, 0.676	0.072, 0.506	-0.035, 0.687	-0.039, 0.687	0.054, 0.588	0.098, 0.302	0.031, 0.595	0.004, 0.966
Interaction Model								
SNP Main Effect	-0.050, 0.807	0.258, 0.155	-0.010, 0.943	0.072, 0.644	0.177, 0.282	0.174, 0.272	-0.064, 0.513	0.048, 0.770
SNP-by-CA Inter	0.293, 0.361	-0.193, 0.497	-0.119, 0.599	0.121, 0.623	0.016, 0.950	0.048, 0.847	0.182, 0.235	0.062, 0.809
SNP-by-AO Inter	0.055, 0846	-0.362, 0.148	0.012, 0.954	-0.366, 0.092	-0.353, 0.120	-0.220, 0.316	0.121, 0.369	-0.159, 0.483
***ERBB2*rs9303274**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	0.032, 0.790	-0.010, 0.926	-0.054, 0.526	-0.010, 0.912	0.042, 0.657	0.030, 0.745	0.008, 0.888	-0.021, 0.826
Interaction Model								
SNP Main Effect	-0.144, 0.475	0.091, 0.616	0.011, 0.935	0.039, 0.804	0.127, 0.416	0.131, 0.402	-0.065, 0.508	0.114, 0.472
SNP-by-CA Inter	0.336, 0.291	-0.127, 0.657	-0.257, 0.247	0.194, 0.432	0.041, 0.867	-0.027, 0.911	0.114, 0.459	-0.130, 0.600
SNP-by-AO Inter	0.231, 0.406	-0.184, 0.465	-0.009, 0.965	-0.249, 0.255	-0.269, 0.215	-0.228, 0.296	0.117, 0.388	-0.252, 0.250
***MIR125A*rs12976445**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.288, 0.016*	0.073, 0.497	0.047, 0.580	-0.223, 0.017*	-0.074, 0.434	0.020, 0.834	-0.023, 0.679	0.038, 0.688
Interaction Model								
SNP Main Effect	-0.619, 0.001*	0.102, 0.565	-0.126, 0.365	-0.456, 0.003*	-0.121, 0.443	-0.017, 0.911	-0.079, 0.391	-0.050, 0.749
SNP-by-CA Inter	0.643, 0.032*	-0.157, 0.573	0.257, 0.238	0.326, 0.168	0.010, 0.968	0.022, 0.927	0.267, 0.065	0.152, 0.532
SNP-by-AO Inter	0.439, 0.099	0.047, 0.849	0.266, 0.171	0.364, 0.085	0.133, 0.544	0.076, 0.723	-0.060, 0.644	0.109, 0.616
***ESR1*rs10484919**	n=203	n=203	n=204	n=203	n=204	n=204	n=204	n=204
Main Effects Model								
SNP Main Effect	-0.037, 0.831	-0.004, 0.976	-0.136, 0.255	0.043, 0.745	-0.221, 0.106	-0.449, 0.001*	-0.107, 0.208	-0.170, 0.202
Interaction Model								
SNP Main Effect	-0.120, 0.655	0.158, 0.502	-0.329, 0.073	0.157, 0.448	-0.054, 0.799	-0.646, 0.001*	-0.214, 0.094	-0.267, 0.194
SNP-by-CA Inter	0.310, 0.569	-0.472, 0.325	0.281, 0.451	-0.105, 0.803	-0.262, 0.542	-0.382, 0.330	-0.211, 0.414	-0.180, 0.666
SNP-by-AO Inter	0.073, 0.846	-0.202, 0.541	0.340, 0.188	-0.199, 0.494	-0.357, 0.231	0.509, 0.062	0.319, 0.075	0.281, 0.330
***ESR1*rs1062577**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	-0.085, 0.610	0.312, 0.035*	-0.143, 0.225	0.058, 0.656	0.004, 0.979	-0.119, 0.360	-0.039, 0.761[b]	-0.119, 0.369
Interaction Model								
SNP Main Effect	-0.023, 0.927	0.380, 0.087	0.131, 0.450	-0.244, 0.211	-0.116, 0.558	-0.018, 0.926	-0.077, 0.525	-0.008,0.970
SNP-by-CA Inter	0.071, 0.869	-0.355, 0.354	-0.696,0.021*	0.646, 0.056	-0.183, 0.594	-0.250, 0.461	-0.074, 0.726	-0.343, 0.317
SNP-by-AO Inter	-0.270, 0.485	0.071, 0.834	-0.352, 0.189	0.378, 0.208	0.599, 0.051	-0.111, 0.714	0.151, 0.420	-0.007, 0.402

***ESR1*rs11964281**	n=213	n=214	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	-0.086, 0.614	-0.031, 0.834	0.089, 0.458	-0.004, 0.974	0.168, 0.218	-0.050, 0.706	0.136, 0.083	-0.014, 0.919
Interaction Model								
SNP Main Effect	0.049, 0.855	-0.167, 0.476	0.282, 0.126	-0.053, 0.798	0.263, 0.216	0.254, 0.217	0.160, 0.194	0.266, 0.203
SNP-by-CA Inter	0.069, 0.869	0.412, 0.264	-0.069, 0.813	0.065, 0.842	-0.084, 0.802	-0.438, 0.177	-0.366, 0.850	-0.358, 0.278
SNP-by-AO Inter	-0.535, 0.190	0.033, 0.927	-0.530, 0.062	0.098, 0.759	-0.300, 0.357	-0.461, 0.145	-0.041, 8.827	-0.487, 0.131
***ESR1*rs12173570**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	-0.025, 0.854	-0.021, 0.866	-0.134, 0.158	0.008, 0.940	-0.051, 0.633	-0.209, 0.048*	-0.001, 0.986	-0.044, 0.684
Interaction Model								
SNP Main Effect	-0.306, 0.154	0.235, 0.221	-0.232, 0.123	0.031, 0.854	0.013, 0.938	-0.269, 0.106	0.012, 0.906	0.003, 0.987
SNP-by-CA Inter	0.408, 0.262	-0.237, 0.467	0.141, 0.580	0.098, 0.732	0.065, 0.821	-0.067, 0.813	-0.251, 0.139	-0.054, 0.850
SNP-by-AO Inter	0.517, 0.091	-0.551, 0.045	0.171, 0.424	-0.116, 0.633	-0.239, 0.327	0.177, 0.454	0.163, 0.251	-0.083, 0.733
***ESR1*rs12665044**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.131, 0.368	-0.040, 0.753	-0.019, 0.850	-0.047, 0.676	0.018, 0.879	-0.049, 0.661	-0.010, 0.887	-0.105, 0.366
Interaction Model								
SNP Main Effect	-0.054, 0.801	-0.057, 0.763	-0.040, 0.789	-0.083, 0.624	0.028, 0.872	0.220, 0.186	-0.025, 0.810	-0.020, 0.905
SNP-by-CA Inter	0.012, 0.972	0.173, 0.578	0.149, 0.546	0.099, 0.718	0.060, 0.831	-0.353, 0.196	0.027, 0.871	-0.110, 0.692
SNP-by-AO Inter	-0.295, 0.387	-0.125, 0.679	-0.067, 0.781	0.013, 0.960	-0.098, 0.716	-0.534, 0.044*	0.031, 0.847	-0.177, 0.512
***ESR1*rs1514348**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.089, 0.502	-0.189, 0.106	-0.002, 0.987	-0.129, 0.212	-0.111, 0.282	0.136, 0.184	0.034, 0.587	0.058, 0.571
Interaction Model								
SNP Main Effect	-0.028, 0.889	-0.040, 0.824	-0.028, 0.845	-0.208, 0.187	-0.070, 0.656	0.139, 0.371	0.063, 0.506	-0.058, 0.715
SNP-by-CA Inter	-0.228, 0.527	-0.161, 0.614	-0.038, 0.881	0.199, 0.479	-0.245, 0.386	-0.026, 0.926	-0.152, 0.372	0.030, 0.915
SNP-by-AO Inter	-0.022, 0.942	-0.321, 0.234	0.097, 0.648	0.093, 0.694	0.024, 0.917	0.003, 0.990	0.015, 0.914	0.292, 0.219
***ESR1*rs1801132**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.151, 0.208	0.117, 0.277	0.069, 0.412	-0.070, 0.459	0.022, 0.820	0.127, 0.175	-0.012, 0.828	0.095, 0.314
Interaction Model								
SNP Main Effect	-0.148, 0.457	0.389, 0.029*	0.095, 0.497	-0.232, 0,141	0.037, 0.814	0.162, 0.297	-0.035, 0.710	-0.096, 0.538
SNP-by-CA Inter	0.246, 0.432	-0.338, 0.228	-0.091, 0.681	0.321, 0.194	0.106, 0.666	0.035, 0.886	-0.009, 0.952	0.136, 0.580

SNP-by-AO Inter	-0.202, 0.468	-0.490, 0.050	-0.008, 0.969	0.184, 0.404	-0.141, 0.521	-0.116, 0.594	0.088, 0.509	0.402, 0.068
***ESR1*rs1884051**	n=206	n=207	n=207	n=206	n=207	n=207	n=207	n=207
Main Effects Model								
SNP Main Effect	-0.212, 0.091	-0.076, 0.485	0.062, 0.479	-0.108, 0.273	-0.048, 0.627	0.153, 0.123	-0.016, 0.782	0.137, 0.172
Interaction Model								
SNP Main Effect	-0.128, 0.526	-0.113, 0.525	0.071, 0.620	-0.198, 0.221	-0.027, 0.861	0.193, 0.230	-0.015, 0.884	0.125, 0.448
SNP-by-CA Inter	0.259, 0.430	0.105, 0.715	-0.095, 0.682	0.272, 0.293	0.322, 0.202	0.017, 0.946	0.923, 0.565	0.022, 0.933
SNP-by-AO Inter	-0.395, 0.164	0.019, 0.938	0.032, 0.871	0.036, 0.872	-0.291, 0.184	-0.113, 0.615	-0.065, 0.640	0.016, 0.945
***ESR1*rs2046210**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.002, 0.987	-0.155, 0.160	0.045, 0.613	0.028, 0.773	0.060, 0.546	-0.008, 0.938	0.078, 0.191	0.124, 0.216
Interaction Model								
SNP Main Effect	-0.012, 0.953	0.014, 0.939	0.041, 0.775	-0.028, 0.859	0.021, 0.895	0.106, 0.503	0.043, 0.666	0.135, 0.404
SNP-by-CA Inter	0.216, 0.501	-0.371, 0.182	0.118, 0.597	0.239, 0.334	0.128, 0.610	-0.076, 0.759	-0.008, 0.957	0.106, 0.675
SNP-by-AO Inter	-0.138, 0.635	-0.152, 0.547	-0.069, 0.735	-0.046, 0.838	0.005, 0.984	-0.242, 0.282	0.094, 0.503	-0.141, 0.538
***ESR1*rs2071454**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.125, 0.409	-0.039, 0.771	-0.075, 0.480	-0.178, 0.128	-0.030, 0.800	0.015, 0.901	0.038, 0.586	-0.041, 0.733
Interaction Model								
SNP Main Effect	-0.087, 0.691	-0.119, 0.539	-0.075, 0.626	-0.288, 0.092	0.042, 0.810	0.221, 0.197	-0.046, 0.652	0.054, 0.759
SNP-by-CA Inter	0.031, 0.933	0.320, 0.328	0.147, 0.574	0.209, 0.467	-0.013, 0.964	-0.300, 0.298	0.131, 0.449	-0.164, 0.577
SNP-by-AO Inter	-0.177, 0.623	-0.044, 0.889	-0.126, 0.618	0.202, 0.468	-0.257, 0.367	-0.393, 0.161	0.197, 0.243	-0.143, 0.617
***ESR1*rs2077647**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	-0.055, 0.698	0.049, 0.698	0.131, 0.188	0.012, 0.914	0.127, 0.259	0.033, 0.769	-0.016, 0.804	0.049, 0.665
Interaction Model								
SNP Main Effect	-0.155, 0.503	-0.069, 0.738	0.227, 0.158	0.225, 0.218	0.057, 0.752	0.231, 0.198	0.071, 0.500	-0.020, 0.912
SNP-by-CA Inter	0.308, 0.418	-0.027, 0.937	-0.080, 0.759	-0.503, 0.088	-0.103, 0.728	-0.172, 0.556	0.174, 0.314	0.343, 0.250
SNP-by-AO Inter	0.089, 0.785	0.347, 0.232	-0.192, 0.395	-0.168, 0.507	0.329, 0.197	-0.400, 0.112	-0.341, 0.022*	-0.055, 0.830
***ESR1*rs2228480**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.083, 0.501	-0.175, 0.107	0.086, 0.316	0.143, 0.138	-0.157, 0.103	0.010, 0.917	0.042, 0.481	0.054, 0.582
Interaction Model								

SNP Main Effect	-0.042, 0.835	-0.079, 0.662	0.097, 0.492	0.087, 0.585	-0.289, 0.068	-0.042, 0.791	-0.001, 0.990	0.116, 0473
SNP-by-CA Inter	0.199, 0.536	-0.153, 0.591	0.186, 0.406	-0.103, 0.682	0.381, 0.127	0.114, 0.652	0.114, 0.461	0.037, 0.884
SNP-by-AO Inter	0.193, 0.501	-0.140, 0.580	-0.135, 0.498	0.214, 0.339	0.081, 0.717	0.056, 0.803	0.024, 0.859	-0.204, 0.369
***ESR1*rs2234693**	n=205	n=205	n=206	n=205	n=206	n=206	n=206	n=206
Main Effects Model								
SNP Main Effect	-0.110, 0.459	0.172, 0.184	0.151, 0.143	0.061, 0.601	0.126, 0.277	0.090, 0.434	-0.043, 0.548	-0.046, 0.698
Interaction Model								
SNP Main Effect	-0.181, 0.454	0.053, 0.797	0.432, 0.008*	0.255, 0.175	0.057, 0.758	0.196, 0.298	0.001, 0.991	-0.118, 0.536
SNP-by-CA Inter	0.192, 0.649	-0.0296, 0.411	-0.520, 0.068	-0.546, 0.097	-0.153, 0.639	-0.203, 0.538	0.213, 0.286	0.386, 0.249
SNP-by-AO Inter	0.084, 0.802	0.485, 0.090	-0.376, 0.096	-0.151, 0.563	0.313, 0.226	-0.160, 0.539	-0.217, 0.171	-0.020, 0.941
***ESR1*rs2347867**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	-0.248, 0.040*	0.000, 0.997	-0.072, 0.396	-0.317, 0.001*	-0.227, 0.016*	-0.087, 0.358	-0.062, 0.280	-0.019, 0.843
Interaction Model								
SNP Main Effect	-0.261, 0.186	0.052, 0.768	-0.155, 0.265	-0.464, 0.002*	-0.299, 0.051	-0.031, 0.839	-0.078, 0.402	-0.078, 0.617
SNP-by-CA Inter	0.076, 0.811	-0.015, 0.960	0.068, 0.760	0.253, 0.296	0.243, 0.321	-0.072, 0.772	-0.010, 0.946	0.000, 1.000
SNP-by-AO Inter	-0.016, 0.955	-0.137, 0.584	0.161, 0.409	0.203, 0.341	0.026, 0.904	-0.095, 0.662	0.055, 0.676	0.157, 0.475
***ESR1*rs2744677**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	-0.163, 0.178	0.087, 0.417	0.015, 0.864	-0.091, 0.337	0.109, 0.250	-0.107, 0.248	-0.009, 0.867	-0.139, 0.140
Interaction Model								
SNP Main Effect	-0.063, 0.751	0.074, 0.675	0.038, 0.783	0.035, 0.825	0.216, 0.167	0.013, 0.930	0.076, 0.417	-0.190, 0.218
SNP-by-CA Inter	0.075, 0.812	-0.140, 0.620	0.251, 0.252	-0.182, 0.464	-0.145, 0.557	-0.011, 0.964	-0.105, 0.483	0.335, 0.172
SNP-by-AO Inter	-0.317, 0.255	0.136, 0.584	-0.197, 0.310	-0.197, 0.371	-0.174, 0.427	-0.299, 0.168	-0.169, 0.201	-0.099, 0.647
***ESR1*rs2813543**	n=212	n=212	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	0.149, 0.250	0.163, 0.153	0.079, 0.374	0.109, 0.281	0.079, 0.445	-0.006, 0.952	-0.030, 0.631	0.009, 0.930
Interaction Model								
SNP Main Effect	-0.087, 0.688	-0.009, 0.961	0.196, 0.185	0.179, 0.293	-0.025, 0.887	-0.136, 0.416	-0.096, 0.356	0.158, 0.347
SNP-by-CA Inter	0.546, 0.103	0.112, 0.702	-0.287, 0.206	-0.054, 0.837	0.207, 0.440	0.195, 0.451	0.099, 0.535	-0.128, 0.621
SNP-by-AO Inter	0.231, 0.461	0.398, 0.146	-0.109, 0.609	-0.151, 0.538	0.129, 0.608	0.202, 0.406	0.107, 0.474	-0.314, 0.198
***ESR1*rs2813544**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								

SNP Main Effect	-0.072, 0.559	0.059, 0.588	-0.045, 0.603	-0.090, 0.349	0.034, 0.731	-0.113, 0.243	-0.013, 0.829	-0.045, 0.648
Interaction Model								
SNP Main Effect	-0.061, 0.759	-0.078, 0.653	-0.058, 0.680	-0.015, 0.925	0.199, 0.212	-0.085, 0.586	0.046, 0.630	-0.059, 0.710
SNP-by-CA Inter	0.343, 0.278	0.433, 0.116	0.090, 0.686	-0.004, 0.986	-0.263, 0.295	0.197, 0.422	0.063, 0.676	0.245, 326
SNP-by-AO Inter	-0.306, 0.286	0.032, 0.898	-0.024, 0.904	-0.195, 0.390	-0.286, 0.213	-0.209, 0.353	-0.273, 0.049*	-0.154, 0.498
***ESR1*rs2941740**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.073, 0.554	0.079, 0.475	-0.052, 0.544	-0.073, 0.451	-0.021, 0.828	-0.009, 0.926	-0.002, 0.975	-0.030, 0.759
Interaction Model								
SNP Main Effect	0.134, 0.513	0.063, 0.732	0.036, 0.799	-0.076, 0.638	0.003, 0.984	-0.318, 0.041*	-0.033, 0.726	-0.258, 0.105
SNP-by-CA Inter	-0.105, 0.753	0.010, 0.972	-0.262, 0.254	0.121, 0.638	0.130, 0.615	0.663, 0.008*	0.294, 0.053	0.554, 0.030*
SNP-by-AO Inter	-0.093, 0.744	0.036, 0.887	-0.069, 0.729	-0.065, 0.770	-0.150, 0.503	0.351, 0.103	-0.112, 0.395	0.210, 0.338
***ESR1*rs3020314**	n=207	n=207	n=208	n=207	n=208	n=208	n=208	n=208
Main Effects Model								
SNP Main Effect	-0.270, 0.034*	-0.055, 0.627	0.011, 0.903	-0.137, 0.168	-0.155, 0.111	0.044, 0.652	-0.074, 0.232	0.115, 0.246
Interaction Model								
SNP Main Effect	-0.172, 0.405	0.011, 0.954	-0.040, 0.783	-0.256, 0.117	-0.226, 0.152	0.025, 0.879	-0.165, 0.100	0.044, 0.786
SNP-by-CA Inter	0.136, 0.695	-0.092, 0.766	-0.024, 0.918	0.212, 0.429	0.311, 0.234	-0.007, 0.979	0.119, 0.473	0.020, 0.941
SNP-by-AO Inter	-0.310, 0.279	-0.124, 0.625	0.127, 0.522	0.161, 0.474	-0.010, 0.963	0.049, 0.825	0.158, 0.252	0.167, 0.455
***ESR1*rs3778099**	n=207	n=208	n=208	n=207	n=208	n=208	n=208	n=208
Main Effects Model								
SNP Main Effect	-0.008, 0.959	-0.139, 0.340	-0.084, 0.457	-0.003, 0.979	0.031, 0.811	0.077, 0.540	0.004, 0.957	-0.021, 0.868
Interaction Model								
SNP Main Effect	0.256, 0.288	-0.006, 0.977	0.005, 0.976	0.128, 0.497	-0.000, 1.000	0.181, 0.333	0.032, 0.777	0.042, 0.827
SNP-by-CA Inter	-0.207, 0.649	-0.060, 0.881	-0.217, 0.487	-0.077, 0.829	0.090, 0.801	-0.039, 0.911	-0.117, 0.589	-0.083, 0.818
SNP-by-AO Inter	-0.690, 0.064	-0.491, 0.133	-0.117, 0.646	-0.323, 0.268	0.035, 0.906	-0.279, 0.334	-0.015, 0.933	-0.116, 0.694
***ESR1*rs3798577**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.042, 0.745	-0.215, 0.056	0.133, 0.137	-0.127, 0.207	-0.149, 0.137	0.117, 0.237	0.083, 0.173	0.078, 0.434
Interaction Model								
SNP Main Effect	-0.074, 0.716	-0.138, 0.443	0.000, 0.997	-0.230, 0.157	-0.306, 0.054	0.044, 0.781	0.153, 0.117	0.058, 0.720
SNP-by-CA Inter	-0.220, 0.495	-0.089, 0.753	0.226, 0.312	0.061, 0.809	0.397, 0.110	0.085, 0.731	-0.170, 0.264	-0.027, 0.915
SNP-by-AO Inter	0.295, 0.331	-0.156, 0.559	0.200, 0.346	0.231, 0.333	0.152, 0.517	0.137, 0.562	-0.060, 0.676	0.071, 0.766

***ESR1*rs488133**	n=208	n=208	n=209	n=208	n=209	n=209	n=209	n=209
Main Effects Model								
SNP Main Effect	-0.038, 0.756	-0.005, 0.963	0.207, 0.015*	0.103, 0.279	0.275, 0.004*	0.143, 0.130	0.117, 0.052	0.277, 0.003*
Interaction Model								
SNP Main Effect	-0.178, 0.378	-0.327, 0.058	0.329, 0.018*	0.279, 0.073	0.182, 0.246	0.365, 0.018*	0.241, 0.012*	0.516, 0.001*
SNP-by-CA Inter	0.343, 0.292	0.200, 0.472	-0.122, 0.581	-0.228, 0.358	0.060, 0.811	-0.198, 0.422	-0.095, 0.535	-0.133, 0.584
SNP-by-AO Inter	0.149, 0.602	0.718, 0.003*	-0.232, 0.234	-0.309, 0.158	0.218, 0.324	-0.425, 0.051	-0.288, 0.034*	-0.544, 0.011*
***ESR1*rs532010**	n=211	n=212	n=212	n=211	n=212	n=212	n=212	n=212
Main Effects Model								
SNP Main Effect	-0.080, 0.546	0.062, 0.591	-0.121, 0.185	-0.073, 0.479	-0.164, 0.121	-0.085, 0.406	-0.054, 0.392	-0.187, 0.073
Interaction Model								
SNP Main Effect	-0.047, 0.815	0.207, 0.237	-0.250, 0.067	-0.163, 0.292	-0.076, 0.635	-0.169, 0.275	-0.098, 0.304	-0.287, 0.067
SNP-by-CA Inter	-0.159, 0.628	-0.190, 0.506	0.420, 0.059	0.116, 0.646	-0.072, 0.782	0.159, 0.527	0.107, 0.491	0.119, 0.640
SNP-by-AO Inter	0.047, 0.884	-0.285, 0.311	0.043, 0.843	0.182, 0.462	-0.240, 0.349	0.119, 0.631	0.044, 0.775	0.217, 0.387
***ESR1*rs6557171**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.234, 0.047*	0.014, 0.898	-0.020, 0.817	-0.273, 0.003*	-0.162, 0.082	-0.091, 0.328	-0.029, 0.616	-0.044, 0.644
Interaction Model								
SNP Main Effect	-0.314, 0.107	0.060, 0.736	-0.043, 0.754	-0.456, 0.003*	-0.167, 0.275	-0.081, 0.599	0.007, 0.940	-0.097, 0.532
SNP-by-CA Inter	0.202, 0.508	0.021, 0.939	-0.006, 0.977	0.322, 0.173	0.073, 0.760	-0.002, 0.993	-0.161, 0.264	-0.109, 0.653
SNP-by-AO Inter	0.072, 0.793	-0.152, 0.541	0.064, 0.741	0.242, 0.252	-0.044, 0.839	-0.025, 0.907	0.031, 0.808	0.228, 0.297
***ESR1*rs77275268**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	0.024, 0.881	-0.141, 0.329	-0.072, 0.532	-0.005, 0.970	-0.208, 0.107	-0.209, 0.098	0.079, 0.295	0.041, 0.751
Interaction Model								
SNP Main Effect	-0.054, 0.850	0.023, 0.926	-0.214, 0.289	0.131, 0.553	-0.050, 0.824	-0.461, 0.035*	-0.032, 0.801	-0.025, 0.914
SNP-by-CA Inter	0.400, 0.426	0.194, 0.660	0.012, 0.973	0.263, 0.496	0.109, 0.782	-0.051, 0.895	0.043, 0.847	-0.031, 0.938
SNP-by-AO Inter	-0.004, 0.992	-0.377, 0.247	0.269, 0.303	-0.338, 0.236	-0.376, 0.198	0.480, 0.089	0.223, 0.175	0.149, 0.611
***ESR1*rs7761133**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	0.044, 0.743	0.000, 0.999	-0.129, 0.170	-0.002, 0.984	0.021, 0.843	-0.057, 0.585	-0.006, 0.924	-0.102, 0.332
Interaction Model								
SNP Main Effect	0.082, 0.690	0.033, 0.856	-0.210, 0.140	-0.101, 0.529	0.065, 0.687	0.163, 0.306	-0.074, 0.449	-0.014, 0.929
SNP-by-CA Inter	0.008, 0.981	-0.022, 0.942	0.166, 0.481	0.195, 0.465	0.042, 0.875	-0.337, 0.203	0.031, 0.851	-0.090, 0.730

SNP-by-AO Inter	-0.139, 0.658	-0.084, 0.766	0.112, 0.606	0.140, 0.571	-0.183, 0.464	-0.362, 0.139	0.190, 0.208	-0.178, 0.461
***ESR1*rs7761846**	n=203	n=203	n=204	n=203	n=204	n=204	n=204	n=204
Main Effects Model								
SNP Main Effect	-0.299, 0.144	-0.158, 0.389	-0.388, 0.005*	-0.316, 0.050	-0.174, 0.284	-0.269, 0.086	-0.148, 0.124	-0.351, 0.022*
Interaction Model								
SNP Main Effect	-0.301, 0.337	0.045, 0.869	-0.344, 0.102	-0.159, 0.530	-0.139, 0.575	-0.254, 0.289	-0.295, 0.041*	-0.274, 0.239
SNP-by-CA Inter	-0.095, 0.840	-0.358, 0.407	0.149, 0.635	-0.233, 0.529	-0.018, 0.963	-0.027, 0.941	0.152, 0.482	-0.140, 0.689
SNP-by-AO Inter	0.097, 0.851	-0.299, 0.510	-0.314, 0.363	-0.373, 0.355	-0.131, 0.749	-0.017, 0.965	0.404, 0.090	-0.122, 0.751
***ESR1*rs7766585**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.104, 0.435	-0.187, 0.122	-0.076, 0.423	-0.052, 0.623	-0.012, 0.913	-0.003, 0.975	0.007, 0.913	-0.009, 0.935
Interaction Model								
SNP Main Effect	-0.020, 0.921	-0.054, 0.770	-0.176, 0.225	-0.009, 0.954	-0.072, 0.660	0.002, 0.988	-0.015, 0.882	-0.104, 0.524
SNP-by-CA Inter	0.124, 0.741	-0.068, 0.841	0.063, 0.814	0.246, 0.403	0.066, 0.825	-0.055, 0.851	0.082, 0.653	0.304, 0.312
SNP-by-AO Inter	-0.288, 0.337	-0.343, 0.208	0.219, 0.305	-0.255, 0.280	0.131, 0.584	0.012, 0.959	-0.001, 0.997	0.066, 0.784
***ESR1*rs7767143**	n=213	n=213	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	-0.008, 0.947	-0.001, 0.994	-0.148, 0.080	-0.042, 0.660	-0.003, 0.978	-0.132, 0.164	-0.067, 0.252	-0.012, 0.901
Interaction Model								
SNP Main Effect	0.012, 0.951	0.275, 0.122	-0.224, 0.099	-0.106, 0.497	0.111, 0.480	0.166, 0.273	-0.194, 0.038*	0.128, 0.407
SNP-by-CA Inter	-0.076, 0.814	-0.322, 0.265	0.219, 0.318	0.153, 0.541	-0.167, 0.514	-0.499, 0.042*	0.088, 0.558	-0.247, 0.323
SNP-by-AO Inter	-0.010, 0.972	-0.523, 0.039*	0.061, 0.750	0.059, 0.790	-0.203, 0.365	-0.419, 0.051	0.310, 0.020*	-0.187, 0.394
***ESR1*rs827421**	n=211	n=211	n=212, 211	n=211	n=212	n=212	n=212	n=212
Main Effects Model								
SNP Main Effect	-0.053, 0.700	0.053, 0.662	0.178, 0.062	0.017, 0.879	0.094, 0.397	0.045, 0.676	-0.019, 0.764	0.046, 0.673
Interaction Model								
SNP Main Effect	-0.233, 0.297	-0.099, 0.617	0.323, 0.035*	0.214, 0.228	0.046, 0.796	0.173, 0.318	-0.011, 0.919	-0.062, 0.724
SNP-by-CA Inter	0.372, 0.319	-0.021, 0.949	-0.159, 0.531	-0.484, 0.099	-0.097, 0.744	-0.114, 0.693	0.259, 0.134	0.379, 0.199
SNP-by-AO Inter	0.249, 0.430	0.422, 0.132	-0.251, 0.244	-0.165, 0.507	0.224, 0.376	-0.257, 0.295	-0.167, 0.255	0.034, 0.890
***ESR1*rs851967**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	0.004, 0.973	0.005, 0.966	-0.102, 0.222	-0.153, 0.099	0.063, 0.508	-0.009, 0.919	-0.068, 0.230	-0.099, 0.291
Interaction Model								

SNP Main Effect	0.024, 0.905	-0.166, 0.345	-0.107, 0.443	-0.075, 0.629	0.073, 0.644	-0.073, 0.632	0.035, 0.717	-0.225, 0.151
SNP-by-CA Inter	-0.056, 0.858	0.207, 0.458	-0.070, 0.749	0.026, 0.914	-0.133, 0.596	0.182, 0.453	-0.067, 0.656	0.274, 0.268
SNP-by-AO Inter	-0.008, 0.977	0.313, 0.208	0.055, 0.781	-0.220, 0.312	0.069, 0.757	0.045, 0.836	-0.230, 0.090	0.128, 0.563
***ESR1*rs851971**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	-0.019, 0.874	-0.005, 0.963	-0.087, 0.296	-0.144, 0.115	0.055, 0.561	-0.027, 0.771	-0.073, 0.187	-0.089, 0.346
Interaction Model								
SNP Main Effect	-0.007, 0.970	-0.163, 0.359	-0.103, 0.459	-0.069, 0.651	0.084, 0.595	-0.115, 0.456	0.025, 0.795	-0.213, 0.176
SNP-by-CA Inter	-0.110, 0.726	0.202, 0.477	-0.024, 0.912	0.018, 0.940	-0.167, 0.504	0.200, 0.412	-0.077, 0.609	0.271, 0.277
SNP-by-AO Inter	0.042, 0.880	0.289, 0.252	0.055, 0.780	-0.214, 0.319	0.045, 0.840	0.091, 0.678	-0.212, 0.116	0.127, 0.569
***ESR1*rs851982**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	0.021, 0.865	0.151, 0.166	-0.027, 0.754	-0.050, 0.602	-0.013, 0.891	-0.059, 0.534	-0.025, 0.661	-0.009, 0.929
Interaction Model								
SNP Main Effect	0.068, 0.736	0.145, 0.422	0.044, 0.759	-0.076, 0.632	-0.012, 0.940	-0.242, 0.123	-0.010, 0.919	-0.203, 0.202
SNP-by-CA Inter	-0.034, 0.919	-0.025, 0.933	-0.234, 0.312	0.270, 0.294	0.066, 0.801	0.377, 0.140	0.218, 0.159	0.442, 0.089
SNP-by-AO Inter	-0.100, 0.722	0.038, 0.880	-0.042, 0.832	-0.099, 0.650	-0.047, 0.834	0.215, 0.320	-0.199, 0.130	0.198, 0.369
***ESR1*rs851998**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	-0.001, 0.991	-0.015, 0.887	-0.088, 0.287	-0.156, 0.086	0.055, 0.554	-0.013, 0.884	-0.061, 0.268	-0.087, 0.348
Interaction Model								
SNP Main Effect	0.000, 1.000	-0.196, 0.267	-0.091. 0.509	-0.084, 0.582	0.065, 0.678	-0.102, 0.504	0.030, 0.748	-0.199, 0.199
SNP-by-CA Inter	-0.052, 0.867	0.239, 0.394	-0.071, 0.744	0.032, 0.894	-0.129, 0.600	0.220, 0.360	-0.056, 0.708	0.265, 0.278
SNP-by-AO Inter	0.028, 0.919	0.316, 0.206	0.051, 0.793	-0.213, 0.321	0.070, 0.752	0.078, 0.715	-0.209, 0.116	0.095, 0.663
***ESR1*rs910416**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.094, 0.489	-0.067, 0.582	0.128, 0.175	0.008, 0.943	0.007, 0.946	0.242, 0.020*	0.001, 0.989	0.070, 0.509
Interaction Model								
SNP Main Effect	-0.084, 0.703	-0.300, 0.121	0.102, 0.509	0.174, 0.323	0.062, 0.723	0.178, 0.288	-0.072, 0.481	0.442, 0.011*
SNP-by-CA Inter	0.176, 0.626	0.085, 0.785	-0.020, 0.934	-0.323, 0.247	-0.028, 0.920	-0.227, 0.398	0.104, 0.529	-0.617, 0.026*
SNP-by-AO Inter	0.351, 0.255	0.589, 0.030*	0.075, 0.729	-0.197, 0.420	-0.125, 0.610	0.286, 0.224	0.111, 0.439	-0.494, 0.042*
***ESR1*rs9322331**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								

SNP Main Effect	-0.105, 0.398	-0.074, 0.516	0.295, 0.777	-0.129, 0.188	-0.111, 0.265	0.036, 0.715	0.052, 0.383	-0.006, 0.955
Interaction Model								
SNP Main Effect	-0.064, 0.744	0.015, 0.934	-0.094, 0.501	-0.205, 0.187	0.071, 0.646	-0.086, 0.576	0.039, 0.682	-0.074, 0.639
SNP-by-CA Inter	-0.019, 0.952	0.107, 0.705	0.260, 0.243	0.152, 0.539	-0.110, 0.657	0.176, 0.476	-0.020, 0.896	-0.028, 0.912
SNP-by-AO Inter	-0.117, 0.692	-0.383, 0.151	0.122, 0.561	0.091, 0.697	-0.475, 0.044*	0.213, 0.361	0.053, 0.712	0.227, 0.340
***ESR1*rs9340799**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.064, 0.618	-0.115, 0.296	0.016, 0.857	-0.107, 0.280	-0.122, 0.228	0.050, 0.612	0.038, 0.530	-0.003, 0.975
Interaction Model								
SNP Main Effect	0.054, 0.784	-0.158, 0.355	-0.081, 0.564	-0.244, 0.111	0.071, 0.646	-0.013, 0.932	0.024, 0.795	-0.049, 0.753
SNP-by-CA Inter	-0.186, 0.565	0.209, 0.449	0.264, 0.245	0.281, 0.257	-0.146, 0.564	0.078, 0.754	0.009, 0.951	-0.048, 0.850
SNP-by-AO Inter	-0.217, 0.476	-0.054, 0.835	0.057, 0.790	0.154, 0.512	-0.489, 0.043*	0.123, 0.604	0.033, 0.820	0.188, 0.439
***ESR1*rs9383938**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.023, 0.889	-0.074, 0.611	-0.094, 0.411	0.060, 0.634	-0.147, 0.248	-0.302, 0.015*	-0.026, 0.735	-0.047, 0.711
Interaction Model								
SNP Main Effect	-0.077, 0.767	-0.025, 0.912	-0.207, 0.256	0.122, 0.543	-0.075, 0.709	-0.458, 0.020*	-0.147, 0.225	-0.095, 0.639
SNP-by-CA Inter	0.435, 0.373	0.236, 0.585	-0.014, 0.966	0.252, 0.503	0.138, 0.717	-0.053, 0.885	0.084, 0.713	0.046, 0.903
SNP-by-AO Inter	0.031, 0.931	-0.209, 0.511	0.246, 0.327	-0.229, 0.409	-0.277, 0.322	0.329, 0.224	0.257, 0.126	0.098, 0.727
***ESR1*rs9397435**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.028, 0.866	-0.195, 0.190	-0.091, 0.44	0.036, 0.786	-0.180, 0.174	-0.257, 0.048*	0.096, 0.212	0.005, 0.969
Interaction Model								
SNP Main Effect	-0.054, 0.850	0.038, 0.880	-0.218, 0.279	0.140, 0.528	-0.038, 0.866	-0.466, 0.033*	-0.039, 0.761	-0.048, 0.831
SNP-by-CA Inter	0.417, 0.407	0.175, 0.690	0.004, 0.990	0.247, 0.527	0.096, 0.808	-0.035, 0.926	0.047, 0.834	-0.001, 0.998
SNP-by-AO Inter	-0.025, 0.947	-0.539, 0.107	0.260, 0.334	-0.304, 0.306	-0.375, 0.213	0.424, 0.145	0.271, 0.112	0.113, 0.707
***ESR1*rs9397456**	n=202	n=202	n=203	n=202	n=203	n=203	n=203	n=203
Main Effects Model								
SNP Main Effect	-0.169, 0.181	0.131, 0.270	0.051, 0.573	-0.045, 0.658	0.003, 0.976	0.095, 0.346	-0.023, 0.705	0.131, 0.206
Interaction Model								
SNP Main Effect	-0.126, 0.540	0.381, 0.049*	0.061, 0.681	-0.200, 0.223	-0.010, 0.952	0.097, 0.554	-0.039, 0.699	-0.071, 0.669
SNP-by-CA Inter	0.118, 0.721	-0.310, 0.313	-0.035, 0.881	0.281, 0.283	0.135, 0.616	0.086, 0.740	-0.021, 0.898	0.118, 0.658
SNP-by-AO Inter	-0.224, 0.446	-0.481, 0.080	-0.003, 0.987	0.217, 0.353	-0.048, 0.840	-0.070, 0.763	0.072, 0.616	0.470, 0.048*

***ESR1*rs985694**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.353, 0.003*	-0.051, 0.647	-0.062, 0.471	-0.140, 0.144	0.022, 0.818	-0.003, 0.978	-0.090, 0.135	0.021, 0.828
Interaction Model								
SNP Main Effect	-0.262, 0.167	-0.119, 0.510	-0.013, 0.927	-0.360, 0.019*	0.002, 0.989	0.080, 0.606	-0.033, 0.740	0.026, 0.871
SNP-by-CA Inter	0.263, 0.392	0.157, 0.590	-0.230, 0.310	0.558, 0.024*	0.124, 0.625	-0.026, 0.919	0.011, 0.947	0.065, 0.798
SNP-by-AO Inter	-0.419, 0.122	0.069, 0.788	0.011, 0.956	0.197, 0.365	-0.044, 0.845	-0.181, 0.413	-0.188, 0.182	-0.058, 0.797
***CCDC170*rs1038304**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.059, 0.663	0.003, 0.982	0.054, 0.570	0.028, 0.793	-0.067, 0.527	0.008, 0.941	-0.087, 0.165	-0.039, 0.716
Interaction Model								
SNP Main Effect	0.215, 0.319	0.012, 0.953	0.027, 0.860	0.226, 0.188	0.081, 0.635	-0.070, 0.683	-0.104, 0.306	-0.149, 0.386
SNP-by-CA Inter	-0.311, 0.383	0.235, 0.471	-0.057, 0.824	-0.197, 0.486	-0.223, 0.428	-0.035, 0.903	0.028, 0.866	-0.035, 0.903
SNP-by-AO Inter	-0.519, 0.085	-0.170, 0.537	0.085, 0.694	-0.372, 0.120	-0.287, 0.228	0.210, 0.382	0.029, 0.841	0.320, 0.182
***CCDC170*rs12662670**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	-0.005, 0.979	-0.363, 0.016*	0.014, 0.910	-0.080, 0.551	-0.164, 0.232	-0.229, 0.087	0.131, 0.093	0.070, 0.608
Interaction Model								
SNP Main Effect	-0.283, 0.339	-0.314, 0.226	-0.086, 0.679	-0.184, 0.426	-0.207. 0.377	-0.435, 0.058	0.089, 0.497	0.016, 0.945
SNP-by-CA Inter	0.613, 0.232	-0.167, 0.709	0.441, 0.223	0.281, 0.483	0.432, 0.288	0.117, 0.767	0.132, 0.564	0.129, 0.751
SNP-by-AO Inter	0.340, 0.385	-0.036, 0.915	0.065, 0.814	0.091, 0.767	-0.068, 0.825	0.390, 0.197	0.071, 0.685	0.068, 0.827
***CCDC170*rs3734805**	n=212	n=212	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	0.057, 0.735	-0.304, 0.042*	-0.001, 0.992	-0.021, 0.876	-0.181, 0.184	-0.201, 0.125	0.106, 0.174	0.039, 0.766
Interaction Model								
SNP Main Effect	-0.085, 0.767	0.025, 0.922	-0.189, 0.347	0.108, 0.630	-0.040, 0.861	-0.455, 0.040*	-0.032, 0.801	-0.023, 0.920
SNP-by-CA Inter	0.447, 0.381	-0.541, 0.233	0.503, 0.159	0.023, 0.954	0.230, 0.574	0.093, 0.811	0.259, 0.256	0.167, 0.677
SNP-by-AO Inter	0.084, 0.826	-0.461, 0.178	0.230, 0.393	-0.259, 0.390	-0.406, 0.190	0.474, 0.109	0.242, 0.160	0.074, 0.807
***CCDC170*rs3757318**	n=212	n=212	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.029, 0.876	-0.334, 0.042*	-0.089, 0.495	0.002, 0.987	-0.115, 0.440	-0.295, 0.043*	0.107, 0.217	-0.005, 0.975
Interaction Model								
SNP Main Effect	-0.218, 0.491	-0.199, 0.471	-0.400, 0.063	0.133, 0.589	-0.237, 0.340	-0.558, 0.022*	-0.052, 0.707	-0.129, 0.597
SNP-by-CA Inter	0.715, 0.210	-0.070, 0.889	0.972, 0.012*	0.060, 0.893	0.699, 0.120	0.475, 0.278	0.521, 0.036*	0.279, 0.527

SNP-by-AO Inter	0.053, 0.902	-0.241, 0.520	0.377, 0.196	-0.268, 0.424	-0.053, 0.876	0.387, 0.243	0.195, 0.297	0.167, 0.615
***CCDC170*rs6929137**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	-0.040, 0.747	-0.322, 0.003*	0.050, 0.566	-0.029, 0.764	-0.061, 0.535	-0.004, 0.969	0.084, 0.162	0.051, 0.596
Interaction Model								
SNP Main Effect	0.099, 0.623	-0.122, 0.479	0.054, 0.697	-0.110, 0.487	-0.081, 0.611	0.046, 0.765	-0.008, 0.932	0.045, 0.772
SNP-by-CA Inter	-0.094, 0.767	-0.428, 0.119	0.033, 0.879	0.217, 0.382	0.031, 0.902	-0.056, 0.816	0.073, 0.624	0.001, 0.997
SNP-by-AO Inter	-0.301, 0.285	-0.201, 0.405	-0.032, 0.871	0.048, 0.828	0.037, 0.869	-0.087, 0.688	0.208, 0.119	0.016, 0.940
***GRB7*rs9910678**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.046, 0.817	-0.094, 0.614	-0.119, 0.404	-0.077, 0.627	-0.031, 0.843	-0.145, 0.355	-0.116, 0.222	-0.429, 0.007*
Interaction Model								
SNP Main Effect	0.162, 0.651	-0.560, 0.085	-0.138, 0.586	0.142, 0.614	-0.009, 0.975	0.143, 0.609	-0.222, 0.194	-0.558, 0.051
SNP-by-CA Inter	-0.482, 0.359	0.491, 0.322	-0.260, 0.482	-0.409, 0.323	-0.334, 0.420	-0.318, 0.438	0.219, 0.382	0.247, 0.554
SNP-by-AO Inter	0.015, 0.976	0.869, 0.048*	0.248, 0.467	-0.235, 0.537	0.194, 0.611	-0.495, 0.190	0.055, 0.811	0.090, 0.814
***GSTM1*rs1065411**	n=208	n=208	n=209	n=208	n=209	n=209	n=209	n=209
Main Effects Model								
SNP Main Effect	0.140, 0.264	0.000, 1.000	-0.101, 0.246	-0.131, 0.172	-0.099, 0.318	0.006, 0.948	-0.048, 0.420	-0.064, 0.511
Interaction Model								
SNP Main Effect	0.218, 0.283	-0.100, 0.581	-0.211, 0.141	-0.275, 0.080	-0.239, 0.140	-0.094, 0.557	0.054, 0.582	-0.069, 0.666
SNP-by-CA Inter	0.129, 0.689	0.217, 0.451	0.222, 0.324	0.301, 0.221	0.312, 0.221	0.249, 0.322	-0.105, 0.500	0.233, 0.354
SNP-by-AO Inter	-0.324, 0.264	0.110, 0.670	0.128, 0.531	0.166, 0.456	0.178, 0.442	0.087, 0.704	-0.231, 0.102	-0.183, 0.422
***GSTM1*rs412543**	n=215	n=215	n=216	n=216	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	-0.058, 0.718	0.012, 0.934	-0.291, 0.009*	-0.124, 0.326	-0.039, 0.763	-0.086, 0.495	-0.188, 0.013*	-0.388, 0.002*
Interaction Model								
SNP Main Effect	0.184, 0.459	-0.041, 0.857	-0.565, 0.001*	-0.220, 0.265	-0.245, 0.230	-0.470, 0.017*	-0.256, 0.032*	-0.901, <0.001*
SNP-by-CA Inter	-0.701, 0.074	0.201, 0.578	0.382, 0.158	0.273, 0.380	0.344, 0.285	0.624, 0.043*	0.126, 0.502	1.012, 0.001*
SNP-by-AO Inter	-0.190, 0.620	-0.055, 0.876	0.513, 0.053	0.031, 0.920	0.313, 0.321	0.640, 0.034*	0.084, 0.645	0.473, 0.116
***NFE2L2*rs35652124**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.063, 0.600	-0.017, 0.876	-0.016, 0.852	-0.280, 0.002*	0.022, 0.814	-0.073, 0.437	0.029, 0.619	0.061, 0.516
Interaction Model								

SNP Main Effect	-0.027, 0.889	-0.264, 0.129	0.172, 0.207	-0.276, 0.066	-0.023, 0.881	-0.248, 0.105	0.038, 0.686	0.255, 0.098
SNP-by-CA Inter	0.150, 0.624	0.371, 0.180	-0.389, 0.072	0.221, 0.347	-0.026, 0.916	0.304, 0.207	-0.075, 0.615	-0.358, 0.142
SNP-by-AO Inter	-0.204, 0.458	0.408, 0.099	-0.235, 0.226	-0.160, 0.450	0.148, 0.497	0.242, 0.262	0.039, 0.772	-0.244, 0.265
***NFE2L2*rs6721961**	n=211	n=212	n=212	n=211	n=212	n=212	n=212	n=212
Main Effects Model								
SNP Main Effect	0.205, 0.183	-0.121, 0.375	-0.081, 0.456	0.194, 0.106	0.063, 0.613	0.097, 0.424	0.078, 0.274	0.082, 0.506
Interaction Model								
SNP Main Effect	0.310, 0.191	-0.014, 0.946	-0.124, 0.459	0.325, 0.077	0.126, 0.504	0.344, 0.063	0.089, 0.415	0.140, 0.459
SNP-by-CA Inter	-0.090, 0.832	0.165, 0.661	-0.005, 0.988	-0.018, 0.956	-0.139, 0.683	-0.466, 0.162	-0.189, 0.336	-0.246, 0.470
SNP-by-AO Inter	-0.230, 0.498	-0.397, 0.187	0.103, 0.667	-0.303, 0.251	-0.085, 0.754	-0.399, 0.134	0.078, 0.621	-0.031, 0.909
***MELK*rs10973007**	n=208	n=209	n=209	n=208	n=209	n=209	n=209	n=209
Main Effects Model								
SNP Main Effect	0.047, 0.717	-0.145, 0.190	0.126, 0.160	0.073, 0.465	0.038, 0.716	0.035, 0.724	0.030, 0.619	0.056, 0.591
Interaction Model								
SNP Main Effect	-0.117, 0.580	0.080, 0.659	-0.215, 0.138	0.085, 0.614	0.123, 0.474	-0.092, 0.572	-0.002, 0.982	-0.131, 0.442
SNP-by-CA Inter	0.067, 0.847	-0.112, 0.701	0.601, 0.010*	-0.028, 0.917	-0.290, 0.292	-0.165, 0.526	-0.282, 0.074	0.028, 0.919
SNP-by-AO Inter	0.386, 0.190	-0.504, 0.046*	0.462, 0.022*	-0.010, 0.964	-0.047, 0.843	0.388, 0.085	0.263, 0.054	0.433, 0.066
***MELK*rs2250340**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.052, 0.764	0.058, 0.705	0.021, 0.865	-0.087, 0.522	0.023, 0.867	0.086, 0.521	0.034, 0.672	-0.085, 0.532
Interaction Model								
SNP Main Effect	-0.168, 0.471	0.001, 0.997	-0.069, 0.671	0.013, 0.945	-0.153, 0.403	0.002, 0.990	-0.067, 0.533	-0.003, 0.988
SNP-by-CA Inter	0.867, 0.069	-0.048, 0.911	0.935, 0.005*	-0.184, 0.626	0.751, 0.045	0.431, 0.247	0.413, 0.063	0.604, 0.101
SNP-by-AO Inter	0.254, 0.524	0.354, 0.329	-0.182, 0.512	-0.191, 0.545	0.101, 0.747	0.000, 1.000	0.069, 0.711	-0.765, 0.014*
***MELK*rs3780350**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.048, 0.725	-0.062, 0.599	-0.170, 0.069	-0.023, 0.826	0.023, 0.832	-0.134, 0.201	-0.018, 0.781	0.001, 0.991
Interaction Model								
SNP Main Effect	-0.048, 0.838	-0.260, 0.211	-0.214, 0.188	-0.306, 0.097	-0.056, 0.767	-0.112, 0.542	-0.024, 0.829	-0.142, 0.450
SNP-by-CA Inter	0.180, 0.620	0.216, 0.496	0.096, 0.698	0.515, 0.067	0.303, 0.296	0.250, 0.372	0.075, 0.658	0.197, 0.495
SNP-by-AO Inter	-0.101, 0.747	0.322, 0.243	0.045, 0.833	0.282, 0.247	-0.002, 0.995	-0.182, 0.453	-0.043, 0.769	0.214, 0.392
***MKI67*rs10732438**	n=210	n=210	n=211	n=211	n=211	n=211	n=211	n=211
Main Effects Model								

SNP Main Effect	-0.030, 0.809	-0.045, 0.679	0.074, 0.393	-0.264, 0.007*	-0.194, 0.047*	0.008, 0.930	-0.038, 0.513	-0.095, 0.330
Interaction Model								
SNP Main Effect	-0.080, 0.699	-0.068, 0.703	0.111, 0.436	-0.371, 0.022*	-0.186, 0.247	0.044, 0.782	-0.080, 0.406	-0.044, 0.786
SNP-by-CA Inter	0.037, 0.907	0.290, 0.295	0.129, 0.559	0.147, 0.555	0.068, 0.785	-0.207, 0.402	0.009, 0.951	-0.022, 0.929
SNP-by-AO Inter	0.108, 0.713	-0.186, 0.464	-0.184, 0.365	0.179, 0.435	-0.062, 0.787	0.050, 0.825	0.111, 0.419	-0.123, 0.594
***MKI67*rs10764751**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.063, 0.597	-0.079, 0.458	-0.017, 0.841	0.037, 0.697	0.046, 0.627	-0.007, 0.940	0.001, 0.984	-0.048, 0.607
Interaction Model								
SNP Main Effect	0.065, 0.743	-0.220, 0.214	-0.086, 0.540	0.142, 0.360	0.041, 0.793	0.140, 0.364	0.075, 0.429	0.140, 0.368
SNP-by-CA Inter	0.173, 0.579	0.223, 0.421	0.087, 0.692	0.099, 0.682	-0.150, 0.541	-0.262, 0.278	-0.141, 0.342	-0.297, 0.223
SNP-by-AO Inter	-0.136, 0.623	0.215, 0.381	0.115, 0.555	-0.331, 0.124	0.120, 0.580	-0.199, 0.356	-0.098, 0.456	-0.273, 0.208
***MMP11*rs131451**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	0.057, 0.699	-0.060, 0.650	-0.074, 0.480	0.140, 0.228	-0.045, 0.699	-0.125, 0.274	-0.006, 0.935	0.047, 0.689
Interaction Model								
SNP Main Effect	0.033, 0.886	-0.157, 0.451	-0.124, 0.449	0.238, 0.189	-0.042, 0.819	-0.237, 0.185	-0.077, 0.498	-0.029, 0.874
SNP-by-CA Inter	0.010, 0.979	0.121, 0.727	-0.057, 0.835	-0.025, 0.932	-0.126, 0.680	0.041, 0.889	-0.012, 0.951	-0.076, 0.801
SNP-by-AO Inter	0.056, 0.870	0.210, 0.494	0.168, 0.487	-0.240, 0.368	0.099, 0.715	0.261, 0.319	0.191, 0.254	0.267, 0.317
***MYBL2*rs11556379**	n=219	n=219	n=220, 219	n=219, 218	n=220, 219	n=220, 219	n=220, 219	n=220, 219
Main Effects Model								
SNP Main Effect	0.027, 0.897	-0.081, 0.655	0.089, 0.532	0.422, 0.008*	0.145, 0.365	0.125, 0.427	0.125, 0.195	-0.016, 0.921
Interaction Model								
SNP Main Effect	-0.208, 0.467	-0.436, 0.082	0.503, 0.010*	0.300, 0.173	0.149, 0.507	0.417, 0.057	0.217, 0.119	0.218, 0.324
SNP-by-CA Inter	0.172, 0.853	1.407, 0.083	-1.568, 0.010*[a]	0.568, 0.510[a]	-0.243, 0.781[a]	0.500, 0.463[a]	-0.109, 0.875[a]	0.306, 0.696[a]
SNP-by-AO Inter	0.524, 0.214	0.695, 0.060	-0.711, 0.014*	0.193, 0.552	0.030, 0.927	-0.581, 0.072	-0.170, 0.407	-0.551, 0.092
***MYBL2*rs2070235**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.311, 0.045*	0.023, 0.872	0.093, 0.401	0.076, 0.537	-0.099, 0.419	-0.356, 0.003*	-0.156, 0.035*	-0.411, 0.001*
Interaction Model								
SNP Main Effect	0.394, 0.101	0.031, 0.886	0.036, 0.832	0.238, 0.214	-0.062, 0.747	-0.602, 0.001*	-0.163, 0.155	-0.663, <0.001*
SNP-by-CA Inter	-0.305, 0.412	-0.191, 0.572	0.056, 0.830	-0.421, 0.156	-0.086, 0.772	0.239, 0.399	0.029, 0.871	0.407, 0.156
SNP-by-AO Inter	0.107, 0.778	0.307, 0.365	0.128, 0.632	-0.081, 0.788	-0.036, 0.904	0.556, 0.055	-0.012, 0.948	0.377, 0.196

***MYBL2*rs619289**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	0.056, 0.664	0.077, 0.508	0.097, 0.278	-0.028, 0.781	-0.034, 0.736	-0.214, 0.031*	-0.051, 0.392	-0.222, 0.028*
Interaction Model								
SNP Main Effect	0.204, 0.325	0.033, 0.857	0.178, 0.219	0.109, 0.498	-0.009, 0.955	-0.250, 0.116	-0.039, 0.689	-0.227, 0.160
SNP-by-CA Inter	-0.201, 0.548	-0.023, 0.939	-0.031, 0.894	-0.293, 0.258	0.065, 0.806	0.001, 0.998	0.019, 0.906	-0.110, 0.670
SNP-by-AO Inter	-0.257, 0.380	0.147, 0.573	-0.188, 0.359	-0.159, 0.484	-0.125, 0.589	0.090, 0.688	-0.056, 0.686	0.107, 0.639
***MYBL2*rs826943**	n=212	n=212	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	-0.023, 0.868	0.014, 0.907	0.130, 0.180	-0.114, 0.296	-0.030, 0.785	-0.099, 0.360	0.059, 0.371	-0.067, 0.532
Interaction Model								
SNP Main Effect	0.002, 0.991	0.016, 0.933	0.255, 0.095	-0.046, 0.788	-0.018, 0.920	-0.132, 0.437	0.026, 0.802	0.011, 0.947
SNP-by-CA Inter	0.117, 0.753	-0.014, 0.967	-0.136, 0.595	-0.124, 0.668	0.121, 0.683	-0.050, 0.861	0.131, 0.462	-0.123, 0.668
SNP-by-AO Inter	-0.150, 0.636	0.003, 0.990	-0.226, 0.301	-0.094, 0.703	-0.121, 0.631	0.113, 0.642	0.002, 0.990	-0.138, 0.573
***MYBL2*rs826944**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	-0.065, 0.635	0.013, 0.915	0.100, 0.301	-0.099, 0.352	-0.002, 0.986	-0.096, 0.365	0.039, 0.541	-0.074, 0.489
Interaction Model								
SNP Main Effect	0.008, 0.938	0.031, 0.873	0.259, 0.091	-0.049, 0.775	-0.017, 0.922	-0.131, 0.438	0.020, 0.847	0.000, 0.999
SNP-by-CA Inter	-0.029, 0.938	-0.119, 0.718	-0.215, 0.403	-0.090, 0.755	0.172, 0.553	-0.052, 0.854	0.066, 0.706	-0.181, 0.530
SNP-by-AO Inter	-0.180, 0.563	0.058, 0.835	-0.267, 0.219	-0.080, 0.742	-0.064, 0.792	0.116, 0.627	0.017, 0.910	-0.083, 0.734
***NDC80*rs12408485**	n=202	n=202	n=203	n=202	n=203	n=203	n=203	n=203
Main Effects Model								
SNP Main Effect	-0.158, 0.200	-0.020, 0.860	-0.059, 0.500	0.049, 0.617	-0.051, 0.601	-0.147, 0.125	-0.056, 0.331	-0.093, 0.350
Interaction Model								
SNP Main Effect	-0.117, 0.577	-0.004, 0.982	0.061, 0.684	-0.290, 0.084	-0.165, 0.322	-0.016, 0.922	-0.103, 0.301	-0.088, 0.603
SNP-by-CA Inter	0.214, 0.536	0.130, 0.686	-0.278, 0.257	0.390, 0.155	0.076, 0.782	-0.230, 0.391	0.115, 0.482	0.104, 0.709
SNP-by-AO Inter	-0.259, 0.363	-0.113, 0.670	-0.123, 0.544	0.589, 0.010*	0.264, 0.246	-0.170, 0.443	0.046, 0.733	-0.082, 0.723
***NDC80*rs2292274**	n=206	n=206	n=207	n=207	n=207	n=207	n=207	n=207
Main Effects Model								
SNP Main Effect	0.117, 0.339	0.051, 0.651	0.114, 0.192	0.019, 0.843	0.114, 0.238	0.095, 0.321	0.054, 0.340	0.130, 0.183
Interaction Model								
SNP Main Effect	0.171, 0.396	0.068, 0.713	0.210, 0.145	0.306, 0.055	0.278, 0.079	0.319, 0.044*	0.041, 0.660	0.219, 0.170
SNP-by-CA Inter	-0.307, 0.333	-0.261, 0.371	-0.127, 0.572	-0.402, 0.107	-0.379, 0.126	-0.322, 0.192	0.054, 0.707	-0.123, 0.623

SNP-by-AO Inter	0.063, 0.822	0.134, 0.605	-0.156, 0.437	-0.447, 0.046*	-0.193, 0.382	-0.349, 0.113	-0.004, 0.973	-0.135, 0.545
***ORC6*rs33994299**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	-0.112, 0.387	-0.048, 0.677	0.103, 0.257	-0.037, 0.715	-0.106, 0.297	0.193, 0.054	0.108, 0.087	0.077, 0.445
Interaction Model								
SNP Main Effect	-0.085, 0.678	-0.162, 0.378	0.158, 0.271	0.110, 0.494	-0.138, 0.393	0.416, 0.008*	0.119, 0.233	0.083, 0.604
SNP-by-CA Inter	-0.041, 0.901	0.169, 0.565	-0.205, 0.367	-0.312, 0.219	0.062, 0.808	-0.417, 0.094	0.001, 0.995	-0.175, 0.488
SNP-by-AO Inter	-0.046, 0.878	0.212, 0.425	-0.010, 0.960	-0.178, 0.442	0.051, 0.828	-0.303, 0.181	-0.043, 0.766	0.131, 0.570
***PGR*rs1042838**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	0.244, 0.075	0.046, 0.712	0.184, 0.058	0.124, 0.253	0.107, 0.330	-0.091, 0.397	0.081, 0.212	0.179, 0.102
Interaction Model								
SNP Main Effect	0.269, 0.207	-0.027, 0.890	0.337, 0.025*	0.087, 0.607	0.016, 0.925	-0.173, 0.302	0.037, 0.712	0.268, 0.118
SNP-by-CA Inter	0.127, 0.717	0.125, 0.694	-0.155, 0.528	0.268, 0.331	0.268, 0.341	0.204, 0.459	0.094, 0.570	-0.039, 0.889
SNP-by-AO Inter	-0.172, 0.590	0.136, 0.640	-0.323, 0.151	-0.103, 0.682	0.072, 0.780	0.065, 0.797	0.058, 0.702	-0.255, 0.319
***PGR*rs1042839**	n=207	n=207	n=208	n=207	n=208	n=208	n=208	n=208
Main Effects Model								
SNP Main Effect	0.192, 0.195	-0.058, 0.655	0.130, 0.211	0.082, 0.477	0.114, 0.331	-0.127, 0.274	0.076, 0.279	0.096, 0.405
Interaction Model								
SNP Main Effect	0.174, 0.433	-0.122, 0.531	0.290, 0.060	0.025, 0.884	0.011, 0.951	-0.187, 0.278	0.037, 0.723	0.263, 0.127
SNP-by-CA Inter	0.257, 0.503	0.224, 0.508	-0.248, 0.352	0.258, 0.389	0.259, 0.390	0.196, 0.511	0.080, 0.658	-0.141, 0.636
SNP-by-AO Inter	-0.128, 0.708	0.016, 0.957	-0.310, 0.193	-0.035, 0.894	0.127, 0.638	0.020, 0.940	0.067, 0.680	-0.433, 0.104
***PGR*rs10895068**	n=213	n=214	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	-0.182, 0.314	-0.277, 0.080	0.053, 0.680	-0.175, 0.211	-0.067, 0.645	0.055, 0.699	-0.079, 0.363	-0.011, 0.937
Interaction Model								
SNP Main Effect	-0.297, 0.294	-0.512, 0.040*	-0.082, 0.679	-0.160, 0.465	0.001, 0.998	-0.136, 0.537	0.114, 0.403	-0.182, 0.425
SNP-by-CA Inter	-0.194, 0.682	0.528, 0.205	-0.011, 0.974	0.171, 0.642	-0.109, 0.775	0.110, 0.767	-0.361, 0.115	0.109, 0.776
SNP-by-AO Inter	0.423, 0.300	0.240, 0.503	0.337, 0.241	-0.174, 0.582	-0.121, 0.714	0.416, 0.192	-0.281, 0.154	0.400, 0.226
***PGR*rs11224561**	n=213	n=213	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	0.052, 0.718	-0.029, 0.817	0.050, 0.618	-0.075, 0.497	0.055, 0.624	0.028, 0.799	0.041, 0.533	-0.137, 0.208
Interaction Model								

SNP Main Effect	-0.058, 0.799	0.012, 0.954	-0.071, 0.653	-0.169, 0.336	0.101, 0.570	-0.028, 0.875	0.184, 0.088	-0.305, 0.072
SNP-by-CA Inter	0.326, 0.410	0.107, 0.757	0.289, 0.294	0.309, 0.308	-0.111, 0.718	-0.075, 0.805	-0.093, 0.618	0.589, 0.045*
SNP-by-AO Inter	0.099, 0.759	-0.195, 0.492	0.162, 0.473	0.052, 0.835	-0.060, 0.811	0.170, 0.493	-0.427, 0.005*	0.032, 0.895
***PGR*rs1893505**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								
SNP Main Effect	0.064, 0.603	0.122, 0.263	-0.043, 0.617	-0.122, 0.207	-0.014, 0.884	-0.071, 0.453	-0.045, 0.443	-0.051, 0.592
Interaction Model								
SNP Main Effect	0.234, 0.260	0.081, 0.660	0.147, 0.307	-0.150, 0.366	-0.175, 0.277	-0.087, 0.589	0.098, 0.316	0.177, 0.269
SNP-by-CA Inter	-0.324, 0.323	-0.060, 0.838	-0.244, 0.285	-0.031, 0.906	0.068, 0.789	0.032, 0.902	-0.294, 0.058	-0.195, 0.440
SNP-by-AO Inter	-0.224, 0.429	0.155, 0.538	-0.295, 0.135	0.089, 0.691	0.399, 0.070	0.018, 0.934	-0.163, 0.221	-0.428, 0.051
***PGR*rs1942836**	n=216	n=216	n=217	n=216	n=217	n=217	n=217	n=217
Main Effects Model								
SNP Main Effect	0.023, 0.851	0.037, 0.741	-0.149, 0.082	-0.049, 0.615	-0.053, 0.590	0.070, 0.473	-0.105, 0.075	-0.175, 0.074
Interaction Model								
SNP Main Effect	0.249, 0.222	0.067, 0.712	-0.071, 0.617	0.157, 0.325	-0.030, 0.852	0.152, 0.342	0.083, 0.355	-0.042, 0.795
SNP-by-CA Inter	-0.538, 0.093	-0.062, 0.830	-0.159, 0.473	-0.383, 0.124	-0.288, 0.252	-0.236, 0.345	-0.440, 0.002*	-0.108, 0.665
SNP-by-AO Inter	-0.219, 0.450	-0.040, 0.878	-0.103, 0.609	-0.276, 0.226	0.157, 0.494	-0.052, 0.821	-0.149, 0.246	-0.288, 0.208
***PGR*rs471767**	n=215	n=215	n=216	n=215	n=216	n=216	n=216	n=216
Main Effects Model								
SNP Main Effect	-0.016, 0.895	0.100, 0.357	-0.022, 0.798	-0.114, 0.224	-0.097, 0.316	0.026, 0.784	-0.011, 0.848	0.026, 0.784
Interaction Model								
SNP Main Effect	-0.134, 0.495	0.132, 0.457	-0.212, 0.124	-0.053, 0.732	0.021, 0.894	0.026, 0.868	-0.088, 0.356	-0.020, 0.898
SNP-by-CA Inter	0.259, 0.405	0.021, 0.941	0.271, 0.210	-0.229, 0.343	-0.076, 0.757	-0.216, 0.374	0.090, 0.548	-0.057, 0.818
SNP-by-AO Inter	0.146, 0.602	-0.116, 0.648	0.318, 0.106	0.001, 0.996	-0.291, 0.193	0.140, 0.526	0.150, 0.270	0.170, 0.449
***PGR*rs474320**	n=196	n=196	n=197	n=196	n=197	n=197	n=197	n=197
Main Effects Model								
SNP Main Effect	0.186, 0.196	0.041, 0.762	0.147, 0.137	-0.009, 0.933	0.066, 0.573	-0.150, 0.167	0.043, 0.518	0.158, 0.163
Interaction Model								
SNP Main Effect	0.248, 0.247	-0.133, 0.506	0.341, 0.022*	-0.079, 0.638	-0.098, 0.574	-0.082, 0.616	0.076, 0.459	0.308, 0.069
SNP-by-CA Inter	0.168, 0.643	0.384, 0.256	-0.392, 0.116	0.352, 0.214	0.296, 0.315	-0.012, 0.965	0.001, 0.994	-0.215, 0.449
SNP-by-AO Inter	-0.356, 0.292	0.225, 0.477	-0.321, 0.168	-0.084, 0.752	0.310, 0.261	-0.214, 0.407	-0.160, 0.321	-0.301, 0.258
***PGR*rs4754732**	n=219	n=219	n=220	n=219	n=220	n=220	n=220	n=220
Main Effects Model								

SNP Main Effect	-0.021, 0.867	0.016, 0.888	0.074, 0.389	0.049, 0.612	0.125, 0.197	0.070, 0.467	0.051, 0.375	0.012, 0.898
Interaction Model								
SNP Main Effect	-0.053, 0.786	0.040, 0.820	-0.015, 0.915	-0.053, 0.732	0.094, 0.533	-0.012, 0.939	-0.042, 0.644	-0.134, 0.380
SNP-by-CA Inter	-0.006, 0.986	0.057, 0.844	0.162, 0.472	0.335, 0.184	0.399, 0.109	0.268, 0.283	0.220, 0.141	0.368, 0.143
SNP-by-AO Inter	0.090, 0.745	-0.105, 0.668	0.130, 0.498	0.055, 0.799	-0.164, 0.441	0.048, 0.821	0.106, 0.405	0.150, 0.486
***PGR*rs484389**	n=211	n=211	n=212	n=211	n=212	n=212	n=212	n=212
Main Effects Model								
SNP Main Effect	0.058, 0.637	0.089, 0.420	-0.005, 0.952	0.063, 0.515	0.047, 0.639	-0.199, 0.037*	-0.066, 0.268	-0.097, 0.323
Interaction Model								
SNP Main Effect	0.112, 0.579	0.014, 0.939	0.318, 0.024*	0.007, 0.967	-0.024, 0.885	-0.220, 0.162	0.012, 0.906	0.160, 0.318
SNP-by-CA Inter	0.240, 0.443	-0.029, 0.919	-0.529, 0.016*	0.191, 0.439	0.069, 0.786	0.128, 0.599	-0.130, 0.401	-0.380, 0.126
SNP-by-AO Inter	-0.377, 0.192	0.270, 0.295	-0.494, 0.015*	0.004, 0.984	0.152, 0.519	-0.044, 0.844	-0.145, 0.313	-0.439, 0.056
***PGR*rs568157**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	0.119, 0.381	0.064, 0.603	0.142, 0.131	0.004, 0.973	0.228, 0.034*	-0.088, 0.405	0.063, 0.320	0.067, 0.527
Interaction Model								
SNP Main Effect	0.057, 0.806	-0.151, 0.468	0.139, 0.390	-0.048, 0.794	0.054, 0.768	-0.383, 0.033*	-0.029, 0.788	-0.055, 0.765
SNP-by-CA Inter	0.052, 0.890	0.353, 0.298	-0.051, 0.842	0.339, 0.241	0.517, 0.074	0.543, 0.057	0.199, 0.250	0.265, 0.363
SNP-by-AO Inter	0.112, 0.715	0.295, 0.286	0.028, 0.896	-0.046, 0.849	0.068, 0.778	0.373, 0.115	0.113, 0.433	0.124, 0.608
***PGR*rs590688**	n=214	n=214	n=215	n=214	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	0.114, 0.404	-0.119, 0.329	0.083, 0.384	-0.072, 0.500	-0.052, 0.632	-0.058, 0.585	0.010, 0.879	0.046, 0.667
Interaction Model								
SNP Main Effect	-0.013, 0.958	-0.222, 0.297	0.228, 0.175	-0.165, 0.381	0.015, 0.937	-0.077, 0.682	0.162, 0.151	0.270, 0.143
SNP-by-CA Inter	0.478, 0.186	0.328, 0.314	-0.235, 0.354	0.430, 0.130	0.217, 0.450	-0.097, 0.731	-0.245, 0.152	-0.194, 0.487
SNP-by-AO Inter	-0.005, 0.988	0.051, 0.858	-0.211, 0.351	-0.050, 0.843	-0.292, 0.256	0.106, 0.676	-0.209, 0.171	-0.416, 0.095
***PGR*rs608995**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	0.063, 0.604	0.051, 0.641	0.011, 0.897	0.055, 0.565	0.057, 0.551	-0.149, 0.115	-0.038, 0.512	-0.069, 0.471
Interaction Model								
SNP Main Effect	0.132, 0.502	0.007, 0.969	0.324, 0.020*	-0.002, 0.990	-0.024, 0.880	-0.167, 0.285	0.015, 0.876	0.192, 0.218
SNP-by-CA Inter	0.176, 0.567	-0.045, 0.872	-0.538, 0.013*	0.167, 0.495	0.129, 0.598	0.132, 0.587	-0.097, 0.517	-0.357, 0.142
SNP-by-AO Inter	-0.369, 0.188	0.176, 0.494	-0.464, 0.019*	0.027, 0.903	0.132, 0.554	-0.052, 0.815	-0.072, 0.599	-0.460, 0.039*

***RACGAP1*rs7303531**	n=213	n=214	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	-0.141, 0.464	-0.197, 0.239	0.057, 0.670	0.160, 0.281	0.017, 0.913	0.119, 0.422	0.056, 0.534	0.223, 0.144
Interaction Model								
SNP Main Effect	-0.546, 0.101	0.015, 0.959	0.023, 0.922	0.368, 0.158	-0.155, 0.562	0.326, 0.207	0.001, 0.997	0.096, 0.718
SNP-by-CA Inter	0.719, 0.155	-0.433, 0.327	0.191, 0.594	-0.241, 0.542	0.524, 0.195	-0.148, 0.704	0.276, 0.240	0.296, 0.461
SNP-by-AO Inter	0.518, 0.233	-0.232, 0.543	-0.011, 0.971	-0.363, 0.286	0.070, 0.840	-0.362, 0.283	-0.113, 0.576	0.124, 0.719
***RFC4*rs1354091**	n=213	n=213	n=214	n=213	n=214	n=214	n=214	n=214
Main Effects Model								
SNP Main Effect	0.047, 0.704	0.174, 0.107	0.057, 0.509	0.080, 0.403	0.173, 0.069	0.077, 0.418	0.094, 0.108	0.014, 0.884
Interaction Model								
SNP Main Effect	-0.068, 0.731	0.165, 0.341	-0.015, 0.915	-0.209, 0.174	0.063, 0.681	0.241, 0.114	0.095, 0.318	0.132, 0.388
SNP-by-CA Inter	0.056, 0.868	0.076, 0.797	0.201, 0.382	0.472, 0.066	0.264, 0.302	-0.290, 0.256	0.071, 0.656	-0.110, 0.667
SNP-by-AO Inter	0.262, 0.352	-0.031, 0.900	0.068, 0.728	0.444, 0.041*	0.133, 0.537	-0.241, 0.262	-0.050, 0.711	-0.234, 0.279
***RRM2*rs1138729**	n=201	n=202	n=202	n=201	n=202	n=202	n=202	n=202
Main Effects Model								
SNP Main Effect	0.123, 0.388	0.030, 0.810	0.148, 0.132	0.008, 0.941	0.017, 0.884	0.083, 0.449	-0.005, 0.939	0.112, 0.336
Interaction Model								
SNP Main Effect	0.256, 0.276	0.151, 0.469	0.245, 0.135	0.323, 0.073	0.137, 0.478	0.048, 0.796	-0.024, 0.835	-0.079, 0.680
SNP-by-CA Inter	-0.002, 0.996	-0.019, 0.953	-0.221, 0.382	-0.207, 0.456	0.046, 0.876	-0.068, 0.812	0.119, 0.500	0.290, 0.327
SNP-by-AO Inter	-0.396, 0.235	-0.312, 0.293	-0.097, 0.677	-0.699, 0.007*	-0.379, 0.166	0.148, 0.572	-0.083, 0.608	0.293, 0.280
***RRM2*rs4309551**	n=217	n=217	n=218	n=217	n=218	n=218	n=218	n=218
Main Effects Model								
SNP Main Effect	-0.024, 0.852	0.021, 0.858	0.114, 0.206	0.035, 0.730	0.106, 0.293	0.102, 0.308	0.047, 0.449	0.048, 0.640
Interaction Model								
SNP Main Effect	0.076, 0.734	0.335, 0.100	0.108, 0.495	0.090, 0.608	0.065, 0.713	0.317, 0.069	-0.083, 0.417	-0.139, 0.437
SNP-by-CA Inter	-0.089, 0.795	-0.466, 0.133	0.183, 0.445	-0.145, 0.590	0.108, 0.687	-0.182, 0.490	0.447, 0.004*	0.341, 0.207
SNP-by-AO Inter	-0.189, 0.538	-0.504, 0.071	-0.094, 0.664	-0.036, 0.881	0.033, 0.893	-0.392, 0.100	-0.001, 0.992	0.216, 0.375
***RRM2*rs4668664**	n=214	n=214	n=215	n=215	n=215	n=215	n=215	n=215
Main Effects Model								
SNP Main Effect	-0.026, 0.829	0.005, 0.964	0.036, 0.677	0.063, 0.506	0.113, 0.234	0.039, 0.676	0.009, 0.871	0.003, 0.978
Interaction Model								
SNP Main Effect	0.080, 0.679	0.073, 0.686	0.009, 0.949	-0.026, 0.867	0.037, 0.814	0.086, 0.582	-0.078, 0.415	-0.042, 0.795
SNP-by-CA Inter	-0.297, 0.343	-0.254, 0.378	0.081, 0.720	-0.010, 0.687	-0.074, 0.768	-0.029, 0.907	0.232, 0.128	-0.010, 0.969

SNP-by-AO Inter	-0.065, 0.813	0.005, 0.983	0.016, 0.937	0.286, 0.194	0.263, 0.235	-0.099, 0.653	0.058, 0.665	0.120, 0.594
***SCUBE2*rs1136966**	n=212	n=213	n=213	n=212	n=213	n=213	n=213	n=213
Main Effects Model								
SNP Main Effect	0.058, 0.646	0.046, 0.673	-0.019, 0.829	0.031, 0.751	0.078, 0.432	0.042, 0.663	0.061, 0.291	-0.019, 0.850
Interaction Model								
SNP Main Effect	0.109, 0.598	-0.037, 0.839	-0.087, 0.544	0.100, 0.538	0.278, 0.092	-0.080, 0.612	0.119, 0.211	-0.189, 0.251
SNP-by-CA Inter	-0.115, 0.726	-0.020, 0.943	-0.003, 0.988	-0.101, 0.691	-0.407, 0.116	0.063, 0.799	-0.204, 0.173	0.227, 0.380
SNP-by-AO Inter	-0.053, 0.855	0.249, 0.326	0.167, 0.407	-0.107, 0.640	-0.279, 0.227	0.260, 0.244	0.000, 0.997	0.284, 0.218
***SCUBE2*rs4910440**	n=218	n=218	n=219	n=218	n=219	n=219	n=219	n=219
Main Effects Model								
SNP Main Effect	-0.020, 0.882	0.081, 0.491	-0.002, 0.987	-0.010, 0.921	-0.046, 0.653	-0.014, 0.892	0.053, 0.388	0.042, 0.686
Interaction Model								
SNP Main Effect	-0.367, 0.083	0.082, 0.666	-0.004, 0.981	-0.238, 0.158	-0.121, 0.462	0.076, 0.641	0.152, 0.116	0.088, 0.593
SNP-by-CA Inter	0.628, 0.059	0.007, 0.983	0.004, 0.985	0.268, 0.306	0.146, 0.573	-0.137, 0.593	-0.083, 0.585	-0.113, 0.664
SNP-by-AO Inter	0.525, 0.088	-0.008, 0.976	0.003, 0.988	0.421, 0.084	0.114, 0.635	-0.153, 0.522	-0.220, 0.119	-0.039, 0.872
***SCUBE2*rs6486125**	n=206	n=206	n=207	n=206	n=207	n=207	n=207	n=207
Main Effects Model								
SNP Main Effect	0.087, 0.476	-0.001, 0.992	-0.116, 0.177	-0.193, 0.043*	-0.095, 0.321	0.042, 0.660	-0.070, 0.238	-0.056, 0.563
Interaction Model								
SNP Main Effect	0.244, 0.215	0.241, 0.175	-0.320, 0.020*	-0.392, 0.011*	0.002, 0.990	-0.043, 0.782	-0.188, 0.051	-0.121, 0.437
SNP-by-CA Inter	-0.653, 0.043*	-0.528, 0.069	0.333, 0.136	0.133, 0.593	-0.187, 0.457	0.065, 0.795	0.181, 0.245	-0.044, 0.863
SNP-by-AO Inter	-0.003, 0.992	-0.282, 0.267	0.313, 0.112	0.417, 0.058	-0.149, 0.501	0.166, 0.453	0.202, 0.141	0.200, 0.369

Note. *=p<0.05; AO=Prescribed Anastrozole Only; CA=Prescribed Chemotherapy plus Anastrozole; Inter=Interaction; SNP=Single Nucleotide Polymorphism.
[a]Coefficient estimation omitted in the robust regression model as the only participant prescribed CA with an *MYBL2*rs11556379 variant allele was removed by the robust regression analysis; results from the standard regression model are reported. [b]Robust regression did not converge after 1000 iterations; results from standard regression model reported. All regression coefficient estimates and p-values reported from robust multiple linear regression models generated using Huber weighting and biweighting iterations. In all models, the prescribed treatment groups, CA and AO, were compared to the reference healthy control group and possession of one or more minor alleles (i.e., homozygous variant genotype plus heterozygous genotype) was compared to the reference, wildtype genotype. All models are adjusted for age, estimated verbal intelligence, levels of depressive symptoms, anxiety, fatigue, and pain, and prescribed treatment group.

Table 19: *ERBB2* SNPs and Cognitive Function Robust Regression Results in Women with Breast Cancer by HER2 IHC Classification Score

Regression coefficient, p-value	Attention	Concentration	Executive Function	Mental Flexibility	Psychomotor Speed	Verbal Memory	Visual Memory	Visual Working Memory
***ERBB2*rs1058808**	n=117	n=117	n=118	n=118	n=118	n=118	n=118	n=118
HER2 IHC Score Effect	-0.184, 0.197	0.010, 0.942	0.088, 0.350	-0.016, 0.878	-0.192, 0.107	-0.065, 0.531	-0.065, 0.330	0.015, 0.888
SNP Effect	0.162, 0.527	0.056, 0.822	-0.162, 0.366	0.227, 0.237	-0.203, 0.367	0.006, 0.979	0.126, 0.322	0.242, 0.220
Interaction Effect	0.147, 0.419	0.006, 0.970	-0.003, 0.983	-0.121, 0.360	0.193, 0.214	0.000, 0.998	-0.055, 0.525	-0.292, 0.033*
***ERBB2*rs1136201**	n=119	n=119	n=120	n=120	n=120	n=120	n=120	n=120
HER2 IHC Score Effect	-0.114, 0.269	-0.040, 0.663	0.067, 0.344	-0.113, 0.133	-0.104, 0.237	-0.032, 0.673	-0.129, 0.010*	-0.190, 0.016*
SNP Effect	-0.377, 0.175	-0.419, 0.108	0.036, 0.853	-0.066, 0.748	-0.038, 0.877	0.133, 0.525	-0.135, 0.320	-0.213, 0.321
Interaction Effect	0.196, 0.296	0.156, 0.108	0.036, 0.782	0.097, 0.482	0.073, 0.656	-0.123, 0.384	0.089, 0.327	0.058, 0.686
***ERBB2*rs1476278**	n=119	n=119	n=120	n=120	n=120	n=120	n=120	n=120
HER2 IHC Score Effect	-0.108, 0.449	0.000, 0.998	0.097, 0.298	0.044, 0.653	-0.172, 0.143	-0.140, 0.167	-0.130, 0.048*	-0.066, 0.526
SNP Effect	0.220, 0.400	-0.083, 0.738	-0.144, 0.420	0.214, 0.258	-0.240, 0.288	-0.074, 0.702	0.067, 0.592	0.201, 0.315
Interaction Effect	0.027, 0.883	0.026, 0.880	-0.010, 0.937	-0.200, 0.122	0.173, 0.259	0.113, 0.393	0.036, 0.673	-0.159, 0.242
***ERBB2*rs1810132**	n=113	n=114	n=114	n=114	n=114	n=114	n=114	n=114
HER2 IHC Score Effect	-0.271, 0.035*	0.028, 0.807	0.098, 0.236	-0.091, 0.327	-0.167, 0.131	-0.063, 0.496	-0.039, 0.534	-0.083, 0.392
SNP Effect	-0.151, 0.570	0.090, 0.714	-0.244, 0.175	-0.053, 0.791	-0.159, 0.506	0.095, 0.636	0.186, 0.167	0.193, 0.359
Interaction Effect	0.321, 0.073	-0.014, 0.930	0.042, 0.724	0.018, 0.890	0.150, 0.343	-0.035, 0.794	-0.095, 0.287	-0.191, 0.171
***ERBB2*rs2517955**	n=119	n=119	n=120	n=120	n=120	n=120	n=120	n=120
HER2 IHC Score Effect	-0.169, 0.268	-0.048, 0.726	0.152, 0.112	-0.010, 0.921	-0.187, 0.128	-0.111, 0.296	-0.062, 0.376	0.018, 0.868
SNP Effect	0.126, 0.630	-0.060, 0.808	-0.157, 0.370	0.171, 0.373	-0.258, 0.252	-0.067, 0.730	0.124, 0.335	0.270, 0.170
Interaction Effect	0.110, 0.558	0.087, 0.610	-0.079, 0.514	-0.117, 0.381	0.184, 0.238	0.067, 0.622	-0.065, 0.466	-0.278, 0.043*
***ERBB2*rs4252596**	n=119	n=119	n=120	n=120	n=120	n=120	n=120	n=120
HER2 IHC Score Effect	-0.127, 0.192	-0.005, 0.959	0.068, 0.331	-0.070, 0.336	-0.037, 0.654	-0.105, 0.147	-0.102, 0.037*	-0.186, 0.016*
SNP Effect	-0.749, 0.020*	-0.143, 0.633	-0.065, 0.774	0.075, 0.750	0.492, 0.070	-0.178, 0.448	-0.066, 0.672	-0.175, 0.483
Interaction Effect	0.387, 0.057	0.069, 0.706	0.032, 0.819	-0.067, 0.642	-0.255, 0.122	0.157, 0.273	0.014, 0.881	0.085, 0.574
***ERBB2*rs903501**	n=112	n=113	n=113	n=113	n=113	n=113	n=113	n=113
HER2 IHC Score Effect	-0.226, 0.084	-0.010, 0.934	0.111, 0.184	-0.043, 0.641	-0.118, 0.298	-0.083, 0.371	-0.089, 0.135	-0.162, 0.101
SNP Effect	-0.132, 0.624	-0.100, 0.690	-0.181, 0.321	0.007, 0.972	-0.151, 0.542	0.058, 0.774	0.125, 0.336	0.135, 0.531
Interaction Effect	0.255, 0.168	0.075, 0.657	-0.002, 0.985	-0.076, 0.575	0.072, 0.665	0.017, 0.901	-0.008, 0.927	-0.067, 0.643

***ERBB2*rs9303274**	n=118	n=118	n=119	n=119	n=119	n=119	n=119	n=119
HER2 IHC Score Effect	-0.059, 0.677	-0.031, 0.813	0.106, 0.255	0.056, 0.573	-0.145, 0.208	-0.142, 0.161	-0.111, 0.100	-0.067, 0.519
SNP Effect	0.274, 0.298	-0.099, 0.696	-0.140, 0.437	0.249, 0.198	-0.184, 0.411	-0.096, 0.623	0.083, 0.525	0.197, 0.332
Interaction Effect	-0.025, 0.892	0.074, 0.665	-0.029, 0.813	-0.215, 0.102	0.144, 0.344	0.113, 0.397	0.008, 0.932	-0.159, 0.249
***MIR125A*rs12976445**	n=119	n=119	n=120	n=120	n=120	n=120	n=120	n=120
HER2 IHC Score Effect	-0.005, 0.966	-0.021, 0.853	-0.011, 0.888	-0.045, 0.611	-0.159, 0.128	0.016, 0.854	-0.101, 0.084	-0.186, 0.049*
SNP Effect	0.097, 0.709	0.002, 0.995	-0.070, 0.691	0.016, 0.933	-0.197, 0.381	0.176, 0.361	-0.037, 0.765	-0.048, 0.810
Interaction Effect	-0.137, 0.430	0.087, 0.584	0.184, 0.114	-0.081, 0.517	0.167, 0.257	-0.173, 0.171	-0.004, 0.964	0.045, 0.732

Note. *=p<0.05; HER2= Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; SNP=Single Nucleotide Polymorphism. All regression coefficient estimates and p-values reported from robust multiple linear regression models generated using Huber weighting and biweighting iterations. In all models, possession of one or more minor alleles (i.e., homozygous variant genotype plus heterozygous genotype) was compared to the reference, wildtype genotype. All models are adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 20: *MKI67* SNPs and Cognitive Function Robust Regression Results in Women with Breast Cancer by Ki67 Index

b-coefficient, p-value	Attention	Concentration	Executive Function	Mental Flexibility	Psychomotor Speed	Verbal Memory	Visual Memory	Visual Working Memory
***MKI67*rs10732438**	n=66	n=65	n=66	n=66	n=65	n=66	n=66	n=66
Ki67 Index Effect	0.002, 0.803	-0.004, 0.528	-0.002, 0.716	-0.015, 0.006*	0.002, 0.844	-0.003, 0.673	-0.002, 0.397	-0.002, 0.717
SNP Effect	0.017, 0.955	-0.120, 0.660	-0.116, 0.597	-0.621, 0.004*	-0.198, 0.425	-0.238, 0.355	-0.135, 0.185	-0.086, 0.705
Interaction Effect	0.008, 0.441	-0.001, 0.940	0.009, 0.221	0.016, 0.023*	0.008, 0.392	0.004, 0.662	0.006, 0.067	0.005, 0.526
***MKI67*rs10764751**	n=68	n=67	n=68	n=68	n=68	n=68	n=68	n=68
Ki67 Index Effect	0.009, 0.102	-0.005, 0.404	0.000, 0.931	-0.002, 0.719	-0.006, 0.220	0.002, 0.630	0.002, 0.486	0.003, 0.584
SNP Effect	0.154, 0.586	-0.005, 0.986	-0.331, 0.158	0.180, 0.459	-0.195, 0.446	0.159, 0.529	0.037, 0.737	0.003, 0.989
Interaction Effect	0.008, 0.455	0.001, 0.958	0.009, 0.325	-0.001, 0.948	0.013, 0.188	-0.010, 0.309	0.000, 0.955	-0.003, 0.733

Note. *=p<0.05; SNP=Single Nucleotide Polymorphism. All regression coefficient estimates and p-values reported from robust multiple linear regression models generated using Huber weighting and biweighting iterations. In all models, possession of one or more minor alleles (i.e., homozygous variant genotype plus heterozygous genotype) was compared to the reference, wildtype genotype. All models are adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 21: Significant (p<0.05) Interaction between Tumor Location Octant and GRS for Visual Working Memory Composite

Visual Working Memory (n=81)	regression coefficient, p-value
Tumor Octant Main Effect	
Lower Outer	0.427, 0.068
Lower Inner	0.081, 0.709
Upper Inner	0.008, 0.969
Upper Middle	0.061, 0.706
Outer Middle	0.180, 0.358
Lower Middle	-4.539, 0.126
Inner Middle	-0.001, 0.998
Visual Working Memory GRS Main Effect	0.437, <0.001*
Interaction	
Lower Outer x GRS	0.479, 0.045*
Lower Inner x GRS	0.243, 0.440
Upper Inner x GRS	-0.174, 0.324
Upper Middle x GRS	-0.292, 0.209
Outer Middle x GRS	0.274, 0.184
Lower Middle x GRS	-2.911, 0.153
Inner Middle x GRS	0.132, 0.793

Note. *=p<0.05; GRS=Genetic Risk/Protection Score. Upper outer octant served as the tumor location octant reference group. Regression coefficient estimates and p-values reported from a robust multiple linear regression model generated using Huber weighting and biweighting iterations. Model adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 22: Significant (p<0.05) Individual CTFs Tested for Interactions with GRSs by Cognitive Function Composite

Cognitive Function Composite	Feature with p<0.05
Attention	Tumor Location Octant
Concentration	Ki67 Classification
Executive Function	-
Mental Flexibility	Tumor Octant Oncotype DX Recurrence Score®
Psychomotor Speed	-
Verbal Memory	Tumor Focality Tumor Laterality PR Status (+/-) HER2 Status (+/-) Oncotype DX Recurrence Score® Squared
Visual Memory	HER2 Status (+/-) HER2 IHC Classification Score Oncotype DX HER2 Score Squared Magee Equation Recurrence Score Squared
Visual Working Memory	Overall TNM Stage Tumor Laterality Tumor Location Octant HER2 Status (+/-) HER2 IHC Classification Score Magee Equation Recurrence Score Squared

Note. CTF=Clinicopathologic Tumor Feature; GRS=Genetic Risk/Protection Score; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors.

Table 23: Statistically Significant (p<0.05) Terms from Individual and Interaction Models Evaluated for Inclusion in Final CTF Model

Cognitive Function Composite	Clinicopathologic Tumor Feature
Attention (n=95)	Tumor Location Octant* Invasive Type ER H-Score* Oncotype DX PR Score* Oncotype DX PR Score Squared Tumor Location Octant by Invasive Type[%] ER H-Score by Oncotype DX PR Score* ER H-Score by Oncotype DX PR Score Squared
Concentration (n=65)	Ki67 Classification* Oncotype DX HER2 Score* Oncotype DX HER2 Score Squared* Oncotype DX Recurrence Score®* Oncotype DX Recurrence Score® Squared* Oncotype DX HER2 Score by Oncotype DX Recurrence Score®* Oncotype DX HER2 Score Squared by Oncotype DX Recurrence Score®* Oncotype DX HER2 Score by Oncotype DX Recurrence Score® Squared* Oncotype DX HER2 Score Squared by Oncotype DX Recurrence Score® Squared*
Executive Function (n=76)	Tumor Location Octant[#] Ki67 Classification[#] Ki67 Index* Oncotype DX ER Score* Tumor Location Octant by Oncotype DX ER Score[#] Tumor Location Octant by Ki67 Classification[#] Tumor Location Octant by Ki67 Index[#] Ki67 Classification by Oncotype DX ER Score[#] Ki67 Index by Oncotype DX ER Score*
Mental Flexibility (n=101)	Tumor Location Octant* Oncotype DX Recurrence Score® Oncotype DX PR Score* Oncotype DX PR Score Squared* Tumor Location Octant by Oncotype DX PR Score* Tumor Location Octant by Oncotype DX PR Score Squared*
Psychomotor Speed (n=75)	Overall TNM Stage Tumor Location Octant* Ki67 Classification[#] Ki67 Index* Oncotype DX PR Score Oncotype DX PR Score Squared Overall TNM Stage by KI67 Classification[#] Tumor Location Octant by KI67 Classification[#%] Ki67 Classification by Oncotype DX PR Score[#] Ki67 Classification by Oncotype DX PR Score Squared[#] Overall TNM Stage by Ki67 Index Overall TNM Stage by Ki67 Index Squared* Tumor Location Octant by Ki67 Index[#] Tumor Location Octant by Ki67 Index Squared[#]
Verbal Memory (n=152)	Tumor Location Octant* Tumor Focality*

	Tumor Laterality PR Status(+/-)* HER2 Status (+/-) Oncotype DX Recurrence Score® Oncotype DX Recurrence Score® Squared PR Status (+/-) by HER2 Status (+/-) Tumor Location Octant by Oncotype DX Recurrence Score®* Tumor Location Octant by Oncotype DX Recurrence Score® Squared*
Visual Memory (n=64)	Overall TNM Stage* HER2 Status (+/-)* HER2 IHC Classification Score Oncotype DX ER Score* Oncotype DX HER2 Score Oncotype DX HER2 Score Squared Magee Equation Recurrence Score* Magee Equation Recurrence Score Squared* HER2 Status (+/-) by Oncotype DX ER Score* Overall TNM Stage by Magee Equation Recurrence Score* Overall TNM Stage by Magee Equation Recurrence Score Squared*
Visual Working Memory (n=133)	Overall TNM Stage Tumor Size* Tumor Size Squared* Tumor Laterality Tumor Location Octant* Nottingham Score* Nottingham Grade# HER2 Status (+/-)# HER2 IHC Classification Score HER2 IHC Classification Score Squared Oncotype DX Recurrence Score® Magee Equation Recurrence Score Magee Equation Recurrence Score Squared Tumor Laterality by Tumor Location Octant* Tumor Location Octant by HER2 Status (+/-)% Tumor Location Octant by Oncotype DX Recurrence Score®* Nottingham Score by HER2 IHC Classification Score* Nottingham Grade by HER2 IHC Classification Score# HER2 IHC Classification Score by Tumor Size* HER2 IHC Classification Score by Tumor Size Squared* Tumor Location Octant by HER2 IHC Classification Score* Tumor Location Octant by HER2 IHC Classification Score Squared# Oncotype DX Recurrence Score® by HER2 IHC Classification Score* Oncotype DX Recurrence Score® by HER2 IHC Classification Score Squared

Note. *=p<0.20, #=omitted from model due to collinearity, %=omitted from model due to limited variability; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors. All term designations obtained from robust multiple linear regression models generated using Huber weighting and biweighting iterations. All models adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

Table 24: R^2 for Final Combined CTF plus GRS and Cognitive Function Regression Models

Cognitive Function Composites	Predictors with p<0.20 from Full CTF Models	R^2 Final CTF Model	R^2 Final CTF plus GRS Model
Attention	Tumor Location Octant ER H-Score Oncotype DX PR Score ER H-Score by Oncotype DX PR Score	0.3075 (n=95)	0.5760 (n=38)
Concentration	Ki67 Classification Oncotype DX HER2 Score Oncotype DX HER2 Score Squared Oncotype DX Recurrence Score® Oncotype DX Recurrence Score® Squared Oncotype DX HER2 Score by Oncotype DX Recurrence Score® Oncotype DX HER2 Score Squared by Oncotype DX Recurrence Score® Oncotype DX HER2 Score by Oncotype DX Recurrence Score® Squared Oncotype DX HER2 Score Squared by Oncotype DX Recurrence Score® Squared	0.2103 (n=66)	0.9548 (n=24)
Executive Function	Ki67 Index Oncotype DX ER Score Ki67 Index by Oncotype DX ER Score	0.3349 (n=76)	0.7286 (n=23)
Mental Flexibility	Tumor Location Octant Oncotype DX PR Score Oncotype DX PR Score Squared Tumor Location Octant by Oncotype DX PR Score Tumor Location Octant by Oncotype DX PR Score Squared	0.5488 (n=101)	0.9456 (n=32)
Psychomotor Speed	Overall TNM Stage Tumor Location Octant Ki67 Index Overall TNM Stage by Ki67 Index Overall TNM Stage by Ki67 Index Squared	0.4230 (n=166)	0.7790 (n=54)
Verbal Memory	Tumor Location Octant Tumor Focality PR Status(+/-) Oncotype DX Recurrence Score® Oncotype DX Recurrence Score® Squared Tumor Location Octant by Oncotype DX Recurrence Score® Tumor Location Octant by Oncotype DX Recurrence Score® Squared	0.4487 (n=158)	0.7932 (n=47)

Visual Memory[a]	Overall TNM Stage HER2 Status (+/-) Oncotype DX ER Score Magee Equation Recurrence Score Magee Equation Recurrence Score Squared HER2 Status (+/-) by Oncotype DX ER Score TNM Stage by Magee Equation Recurrence Score TNM Stage by Magee Equation Recurrence Score Squared	0.3241 (n=94)	0.7249 (n=31)
Visual Working Memory[b]	Tumor Size Tumor Size Squared Tumor Laterality Tumor Location Octant Nottingham Score HER2 IHC Classification Score Oncotype DX Recurrence Score® Tumor Laterality by Tumor Location Octant Tumor Location Octant by Oncotype DX Recurrence Score® Nottingham Score by HER2 IHC Classification Score HER2 IHC Classification Score by Tumor Size HER2 IHC Classification Score by Tumor Size Squared Tumor Location Octant by HER2 IHC Classification Score Oncotype DX Recurrence Score® by HER2 IHC Classification Score	0.5068 (n=137)	0.9531 (n=44)

Note. CTF=Clinicopathologic Tumor Feature; GRS=Genetic Risk/Protection Score; ER=Estrogen Receptor; HER2=Human Epidermal Growth Factor Receptor 2; IHC=Immunohistochemistry; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay; PR=Progesterone Receptor; TNM=Tumor, Node, Metastasis Classification of Malignant Tumors. [a]In the combined CTF plus GRS regression model for visual memory, the interactions between HER2 status and Oncotype DX ER Score, TNM Stage 2b and Magee Equation Recurrence Score, and TNM Stage 2b and Magee Equation Recurrence Score Squared were omitted due to collinearity. [b]In the combined CTF plus GRS regression model for visual working memory, the interactions between upper middle, inner middle, and lower middle tumor location octants and HER2 IHC classification score were omitted due to collinearity. The statistically significant interaction between visual working memory GRS and tumor location octant identified from the CTF-by-GRS interaction analysis was also omitted from the model due to collinearity. R^2 obtained from standard multiple linear regression models. All models are adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain.

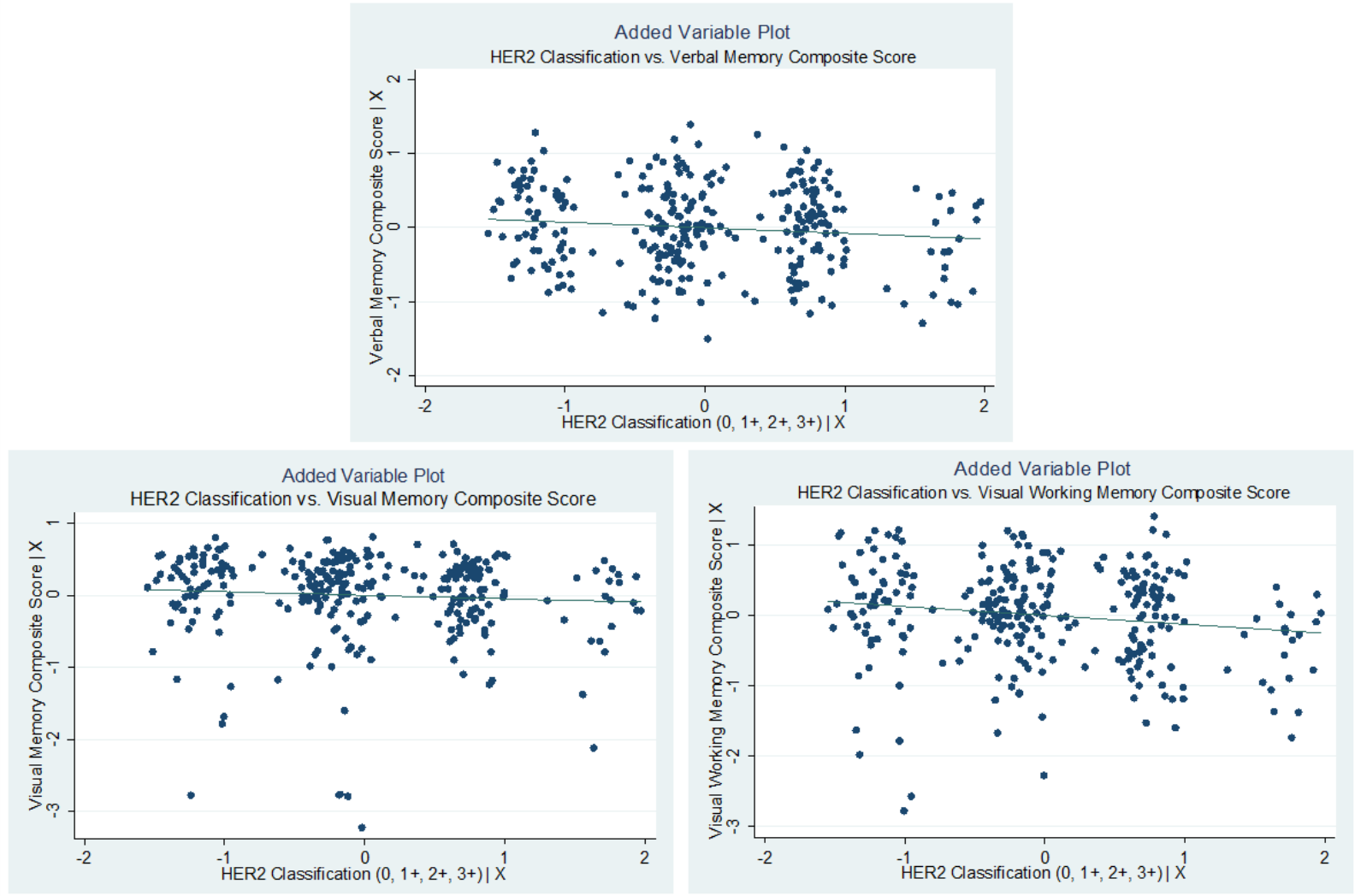


Figure 4: HER2 IHC Classification Score vs. Verbal, Visual, and Visual Working Memory Composite Score Added Variable Plots

Note. X=age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, pain; HER2=Human Epidermal Growth Factor Receptor 2. Figure generated using Stata® Data Analysis and Statistical Software SE Version 14.1 (StataCorp, College Station, TX).

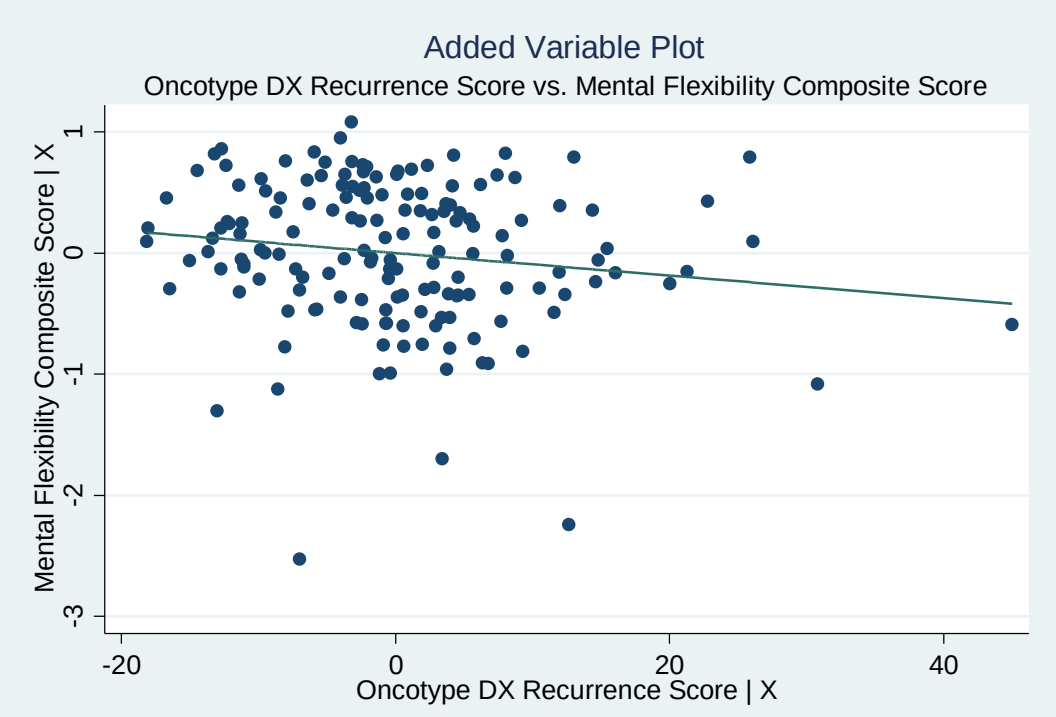
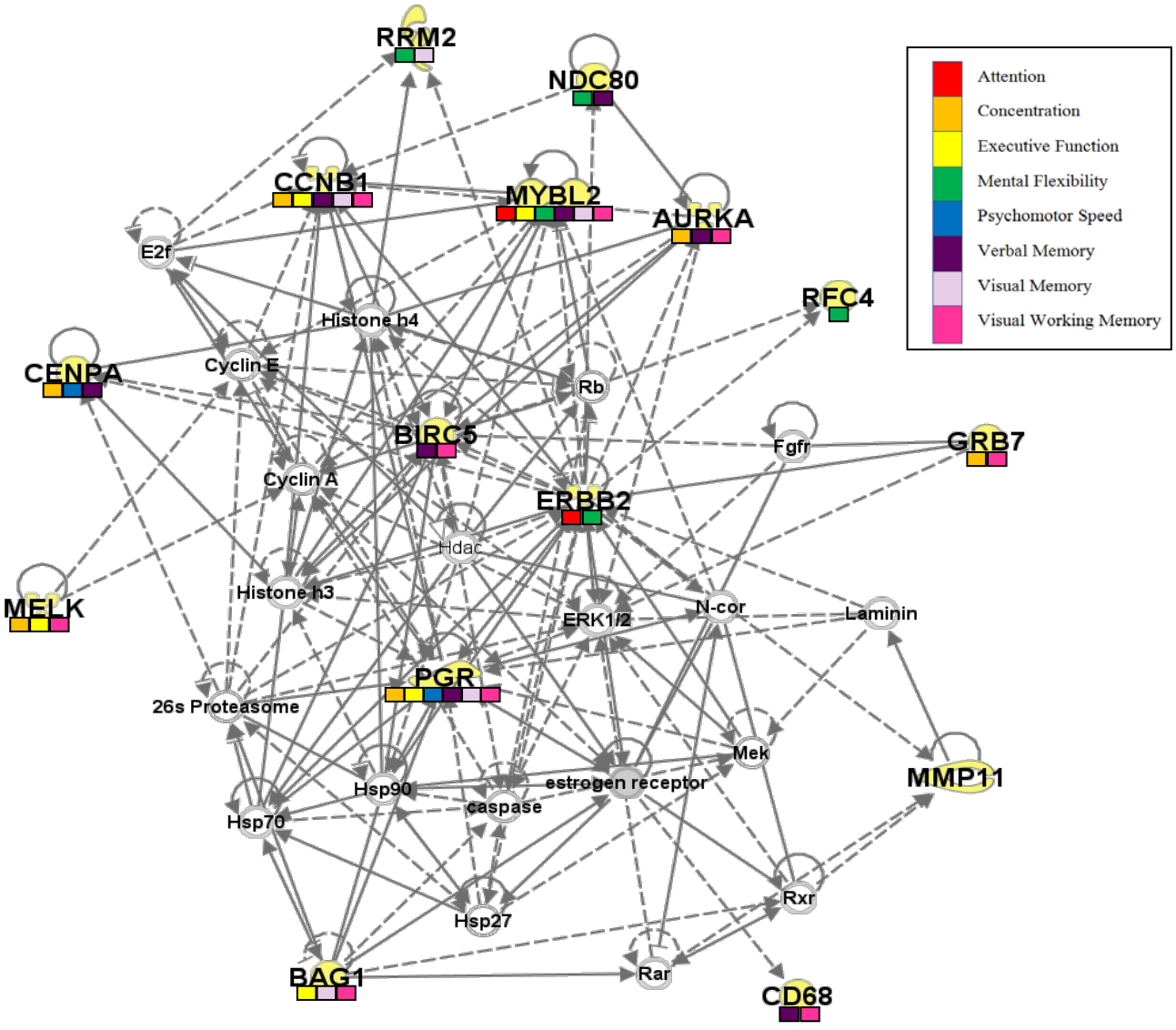


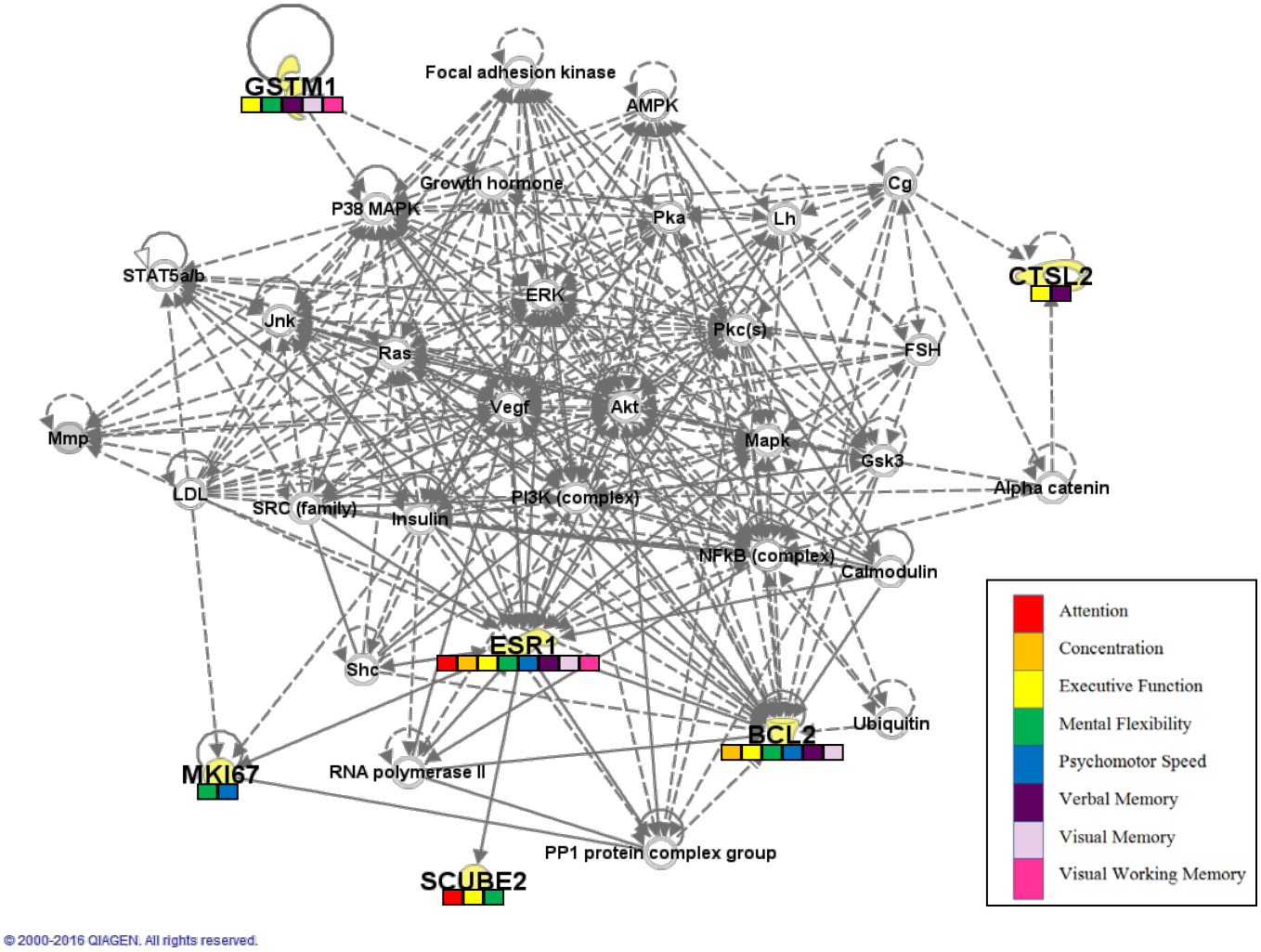
Figure 5: Oncotype DX® Breast Cancer Assay Recurrence Score vs. Mental Flexibility Composite Score Added Variable Plot

Note. X=age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, pain; Oncotype DX=Genomic Health Inc. Oncotype DX® Breast Cancer Assay. Figure generated using Stata® Data Analysis and Statistical Software SE Version 14.1 (StataCorp, College Station, TX).

a.



b.



c.

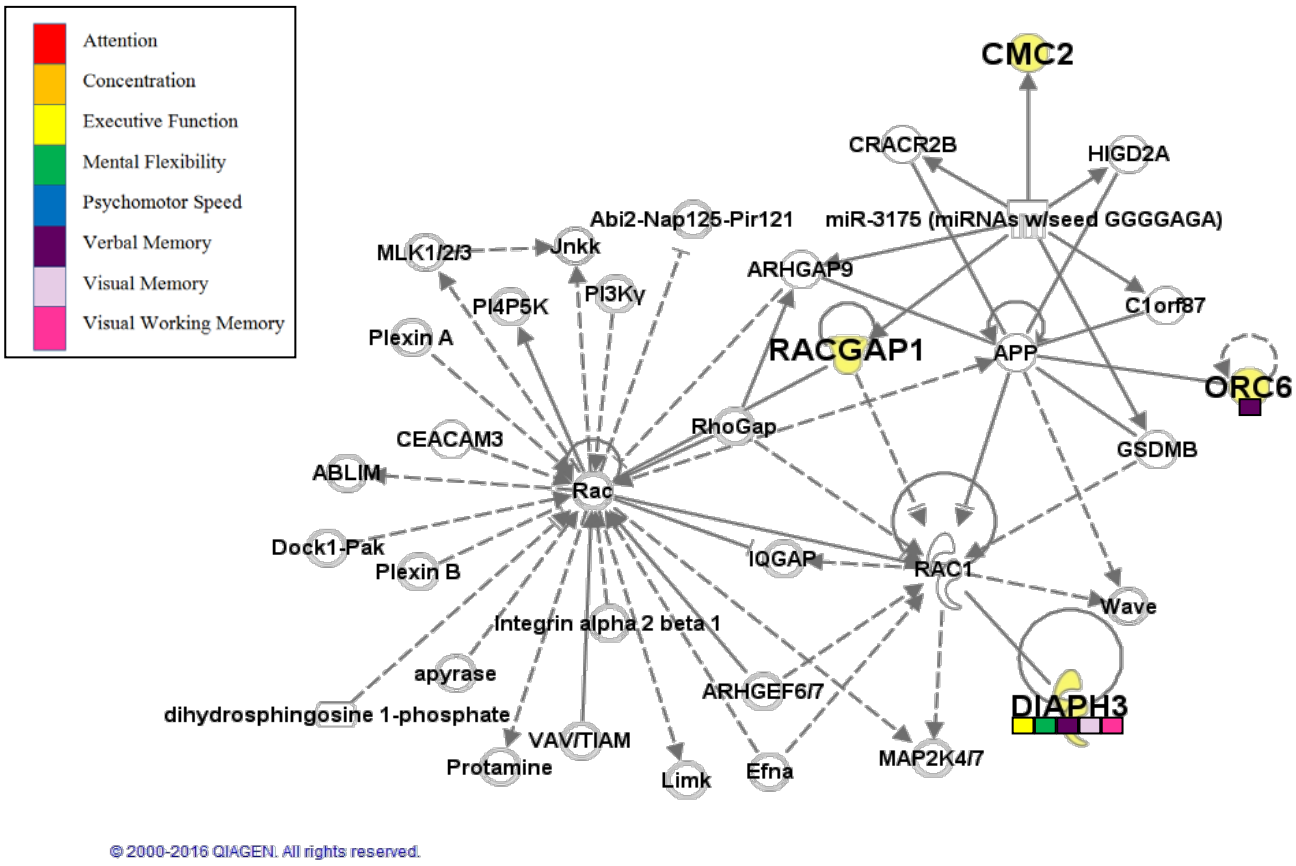


Figure 6: Candidate Gene-Gene Networks Generated Using QIAGEN's Ingenuity Pathway Analysis

Note. The networks were generated through use of QIAGEN's Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City, www.quigen.com/ingenuity). Candidate genes are highlighted in yellow. Colored blocks correspond to cognitive function composites that were found to have one or more significant SNP main effects and/or SNP-by-prescribed treatment group interactions for a particular gene. The main associated functions of each network are as follows: (a) cancer, organismal injury and abnormalities, and reproductive system disease; (b) cancer, cellular development, and cellular growth and proliferation; and (c) cellular assembly and organization, cell morphology, and cellular function and maintenance.

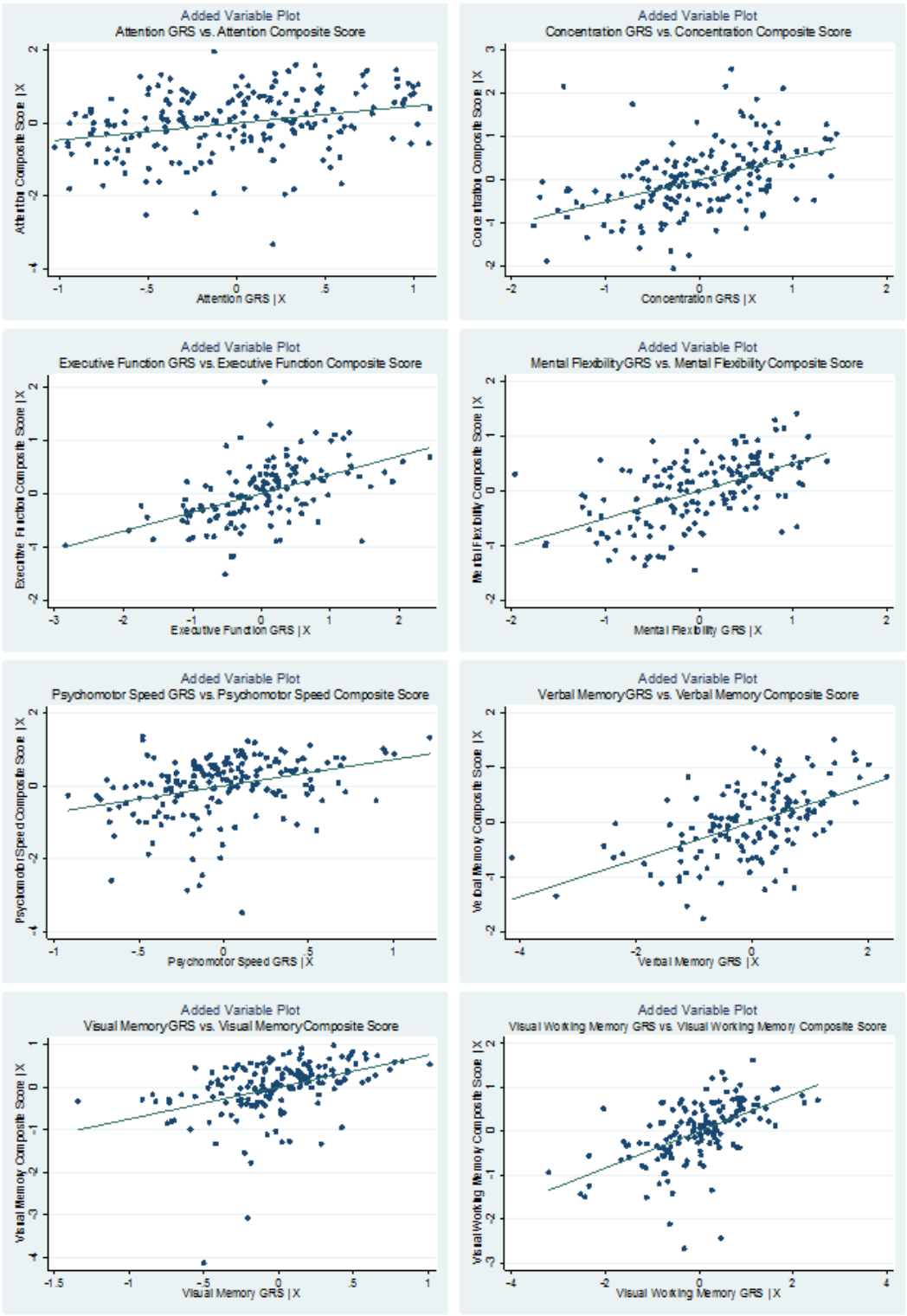


Figure 7: GRS by Cognitive Function Composite Score Added Variable Plots

Note. X=age, estimated verbal intelligence, levels of depressive symptoms, anxiety, fatigue, pain, and prescribed treatment; GRS=Genetic Risk/Protection Score. Figure generated using Stata® Data Analysis and Statistical Software SE Version 14.1 (StataCorp, College Station, TX).

APPENDIX C

UNIVERSITY OF PITTSBURGH INSTITUTIONAL REVIEW BOARD

APPROVAL LETTERS

University of Pittsburgh
Institutional Review Board

3500 Fifth Avenue
Pittsburgh, PA 15213
(412) 383-1480
(412) 383-1508 (fax)
http://www.irb.pitt.edu

Memorandum

To: Theresa Timcheck-Koleck, PhD, BSN
From: Sue Beers, PhD, Vice Chair
Date: 6/10/2013
IRB#: PRO13040672
Subject: Cognitive Function and Breast Cancer: Genomics and Disease Characteristics

---

The University of Pittsburgh Institutional Review Board reviewed and approved the above referenced study by the expedited review procedure authorized under 45 CFR 46.110. Your research study was approved under 45 CFR 46.110 (3) (7).

This study is supported by the following federal grant application:
Anticipated Cognitive Function and Breast Cancer: Genomics and Disease Characteristics

The risk level designation is Minimal Risk.

Approval Date: 6/7/2013
Expiration Date: 6/6/2014

For studies being conducted in UPMC facilities, no clinical activities can be undertaken by investigators until they have received approval from the UPMC Fiscal Review Office.

Please note that it is the investigator's responsibility to report to the IRB any unanticipated problems involving risks to subjects or others [see 45 CFR 46.103(b)(5)]. Refer to the IRB Policy and Procedure Manual regarding the reporting requirements for unanticipated problems which include, but are not limited to, adverse events. If you have any questions about this process, please contact the Adverse Events Coordinator at 412-383-1480.

The protocol and consent forms, along with a brief progress report must be resubmitted at least one month prior to the renewal date noted above as required by FWA00006790 (University of Pittsburgh), FWA00006735 (University of Pittsburgh Medical Center), FWA00000600 (Children's Hospital of Pittsburgh), FWA00003567 (Magee-Womens Health Corporation), FWA00003338 (University of Pittsburgh Medical Center Cancer Institute).

Please be advised that your research study may be audited periodically by the University of Pittsburgh Research Conduct and Compliance Office.

University of Pittsburgh
Institutional Review Board

3500 Fifth Avenue
Pittsburgh, PA 15213
(412) 383-1480
(412) 383-1508 (fax)
http://www.irb.pitt.edu

Memorandum

To: Theresa Koleck
From: Christopher Ryan PhD, Vice Chair
Date: 4/8/2014
IRB#: REN14030156 / PRO13040672
Subject: Cognitive Function and Breast Cancer: Genomics and Disease Characteristics

---

Your renewal for the above referenced research study has received expedited review and approval from the Institutional Review Board under:
45 CFR 46.110.(3)
45 CFR 46.110.(7)

Please note the following information:

Approval Date: 4/8/2014
Expiration Date: 4/7/2015

Please note that it is the investigator's responsibility to report to the IRB any unanticipated problems involving risks to subjects or others [see 45 CFR 46.103(b)(5) and 21 CFR 56.108(b)]. Refer to the IRB Policy and Procedure Manual regarding the reporting requirements for unanticipated problems which include, but are not limited to, adverse events. If you have any questions about this process, please contact the Adverse Events Coordinator at 412-383-1480.

The protocol and consent forms, along with a brief progress report must be resubmitted at least **one month** prior to the renewal date noted above as required by FWA00006790 (University of Pittsburgh), FWA00006735 (University of Pittsburgh Medical Center), FWA00000600 (Children's Hospital of Pittsburgh), FWA00003567 (Magee-Womens Health Corporation), FWA00003338 (University of Pittsburgh Medical Center Cancer Institute).

Please be advised that your research study may be audited periodically by the University of Pittsburgh Research Conduct and Compliance Office.

University of Pittsburgh
Institutional Review Board

3500 Fifth Avenue
Pittsburgh, PA 15213
(412) 383-1480
(412) 383-1508 (fax)
http://www.irb.pitt.edu

Memorandum

To: Theresa Koleck
From: IRB Office
Date: 4/3/2015
IRB#: REN15020164 / PRO13040672
Subject: Cognitive Function and Breast Cancer: Genomics and Disease Characteristics

---

Your renewal for the above referenced research study has received expedited review and approval from the Institutional Review Board under 45 CFR 46.110 (3) (7).

Please note the following information:

Approval Date: 4/3/2015
Expiration Date: 4/2/2016

Please note that it is the investigator's responsibility to report to the IRB any unanticipated problems involving risks to subjects or others [see 45 CFR 46.103(b)(5) and 21 CFR 56.108(b)]. Refer to the IRB Policy and Procedure Manual regarding the reporting requirements for unanticipated problems which include, but are not limited to, adverse events. If you have any questions about this process, please contact the Adverse Events Coordinator at 412-383-1480.

The protocol and consent forms, along with a brief progress report must be resubmitted at least **one month** prior to the renewal date noted above as required by FWA00006790 (University of Pittsburgh), FWA00006735 (University of Pittsburgh Medical Center), FWA00000600 (Children's Hospital of Pittsburgh), FWA00003567 (Magee-Womens Health Corporation), FWA00003338 (University of Pittsburgh Medical Center Cancer Institute).

Please be advised that your research study may be audited periodically by the University of Pittsburgh Research Conduct and Compliance Office.

University of Pittsburgh
Institutional Review Board

3500 Fifth Avenue
Pittsburgh, PA 15213
(412) 383-1480
(412) 383-1508 (fax)
http://www.irb.pitt.edu

Memorandum

To: Theresa Koleck
From: IRB Office
Date: 3/29/2016
IRB#: REN16020165 / PRO13040672
Subject: Cognitive Function and Breast Cancer: Genomics and Disease Characteristics

---

Your renewal for the above referenced research study has received expedited review and approval from the Institutional Review Board under: 45 CFR 46.110 (3) (7).

Please note the following information:

Approval Date: 3/29/2016
Expiration Date: 3/28/2017

Please note that it is the investigator's responsibility to report to the IRB any unanticipated problems involving risks to subjects or others [see 45 CFR 46.103(b)(5) and 21 CFR 56.108(b)]. Refer to the IRB Policy and Procedure Manual regarding the reporting requirements for unanticipated problems which include, but are not limited to, adverse events. If you have any questions about this process, please contact the Adverse Events Coordinator at 412-383-1480.

The protocol and consent forms, along with a brief progress report must be resubmitted at least **one month** prior to the renewal date noted above as required by FWA00006790 (University of Pittsburgh), FWA00006735 (University of Pittsburgh Medical Center), FWA00000600 (Children's Hospital of Pittsburgh), FWA00003567 (Magee-Womens Health Corporation), FWA00003338 (University of Pittsburgh Medical Center Cancer Institute).

Please be advised that your research study may be audited periodically by the University of Pittsburgh Research Conduct and Compliance Office.

APPENDIX D

MANUSCRIPT #1: MOLECULAR GENOMIC RESEARCH DESIGNS

Published in final edited form as:
Annu Rev Nurs Res. 2011 ; 29: 1–26.

Molecular Genomic Research Designs

Kelley Baumgartel[1], **Jamie Zelazny**[1], **Theresa Timcheck**[1], **Chantel Snyder**[1], **Mandy Bell**[1], and **Yvette Conley**[1,2,*]

[1]University of Pittsburgh School of Nursing

[2]University of Pittsburgh Department of Human Genetics

Abstract

Genetic and genomic research approaches have the capability to expand our understanding of the complex pathophysiology of disease susceptibility, susceptibility to complications related to disease, trajectory of recovery from acquired injuries and infections, patient response to interventions and therapeutics, as well as informing diagnoses and prognoses. Nurse scientists are actively involved in all of these fields of inquiry and the goal of this manuscript is to assist with incorporation of genetic and genomic trajectories into their research and facilitate the design and execution of these studies. New studies that are going to embark on recruitment, phenotyping, and sample collection will benefit from forethought about research design to ensure that it addresses the research questions or hypotheses being tested. Studies that will utilize existing data or samples will also benefit from forethought about research design for the same reason but to also address the fact that some designs may not be feasible with the available data or samples. This manuscript discusses candidate gene association, genome wide association, candidate gene expression, global gene expression, and epigenetic/epigenomic study designs. Information provided includes rationale for selecting an appropriate study design, important methodology considerations for each design, key technologies available to accomplish each type of study, and online resources available to assist in executing each type of study design.

In the last decade we have progressed from a rough draft of the human genome sequence to availability of an abundance of publicly available databases and high throughput data collection technologies to facilitate genetic and genomic study design. Genetic (focus on one gene at a time) and genomic (focus on entire genome as well as gene-gene interactions) research continues to hold great promise for understanding a wealth of human conditions, providing objective data for diagnosis and prognosis, informing therapeutics, and providing the cornerstone for evidence based practice for genomic health care (Green, Guyer, & National Human Genome Research Institute [NHGRI], 2011; Lander, 2011). The research programs of many nurse scientists are ripe for incorporating a genetic/genomic research component or movement of existing genetic or genomic research in a new direction.

The goal of his paper is to bring together key information about designing studies with a molecular genetic or genomic focus coupled with dynamic resources offered to the reader to

*Corresponding Author: Yvette P. Conley, University of Pittsburgh, 3500 Victoria Street, 440 Victoria Building, Pittsburgh, PA 15261, yconley@pitt.edu, 412-383-7641.

Contributor's list:
Kelley Baumgartel, RN, Doctoral Student, University of Pittsburgh, Pittsburgh, PA
Jamie Zelazny, RN, MPH, Doctoral Student, University of Pittsburgh, Pittsburgh, PA
Theresa Timcheck, RN, Doctoral Student, University of Pittsburgh, Pittsburgh, PA
Chantel Snyder, RN, Doctoral Student, University of Pittsburgh, Pittsburgh, PA
Mandy Bell, RN, Doctoral Student, University of Pittsburgh, Pittsburgh, PA
Yvette P. Conley, PhD, Associate Professor of Nursing and Human Genetics, University of Pittsburgh, Pittsburgh, PA

expand their understanding and ensure access to state of the science information. It is not meant to be an exhaustive resource, but one that sets the stage for contemplation of embarking on such research designs and key issues to ponder during study design phase. This paper is written for the researcher who has a basic understanding of genetics and is contemplating adding a genetic or genomic component to their research or designing the next step in their genetic or genomic program of research. Readers are encouraged to visit an extremely useful resource, the National Human Genome Research Institute's talking glossary at http://www.genome.gov/glossary, for clarification of unfamiliar terms and expansion of knowledge about genetic terminology. Technology to collect genetic and genomic data changes rapidly, therefore proper study design, and selection of appropriate methodology to accomplish a study also change rapidly. This paper incorporates a large number of online resources that are continuously updated in an attempt to keep the paper as up to data as possible. Readers are encouraged to visit these online resources when designing their study to ensure that their study design is state of the science.

DNA POLYMORPHISM BASED ASSOCIATION STUDIES

The overall objective of a polymorphism based association study is to examine the relationship between DNA variation and a phenotype (e.g., diabetes, fatigue). A polymorphism is defined as a DNA variation that is present in at least one percent of the population (NHGRI, n.d.). One advantage of this approach compared to other genetic/genomic approaches is the use of DNA. DNA is a very stable template for experiments, allowing for use of previously collected samples. Such a retrospective approach could save time and money that would be needed to prospectively recruit participants and collect samples; however, attention must be given to subject consent to assure that informed consent was obtained for future genetic/genomic evaluation related to the phenotype of interest. Another advantage is that this approach does not require that subjects be related, which is a requirement for linkage analysis, an approach not discussed in this manuscript. It should be noted that while related individuals are not required, newer software has been developed to allow for the analyses of related individuals within the context of an association study. Two very appealing additional advantages of polymorphism based studies are the fact that polymorphisms do not change over time and the DNA template that is utilized can be extracted from any tissue. The sample for DNA extraction and collection of polymorphism data only need to be collected once, yet that polymorphism data can be evaluated within the context of a phenotype that changes over time. While blood and saliva are the most frequently used cell/tissue type for DNA extraction, any cells/tissues that have a nucleus can serve as samples for polymorphism based studies. Because DNA polymorphisms do not change and are not tissue specific, investigators need not worry about collection of DNA samples over time or from what tissue DNA extraction occurs. These advantages are not carried over to other genomic approaches detailed in this manuscript.

Candidate Gene Association Studies

Rationale for taking a candidate gene association approach—Candidate gene association studies investigate polymorphisms representing a specific gene(s) to determine if it is associated with a phenotype of interest. With this hypothesis-driven approach, the investigator pre-selects the candidate gene(s) to be evaluated. This approach is only appropriate when *a priori* assumptions about the gene(s) that may be involved in the phenotype of interest can be justified.

Genome wide association studies (GWAS); discussed in the next section, have large sample size requirements (e.g., 1000 cases/1000 controls), and one relative advantage of the candidate gene approach is that it often requires half that number or less. This reduced sample size requirement compared to a GWAS is due to the focused evaluation of a

candidate gene(s), which reduces multiple testing concerns. The candidate gene association approach is also ideal when studying rarer phenotypes since attainment of a large sample may not be feasible for a condition with a low population frequency.

Subject and sample considerations—Clearly defined inclusion/exclusion criteria, which include a detailed definition of the phenotype, are essential to the candidate gene association approach. Structured inclusion/exclusion criteria help to ensure that individuals with/without the phenotype of interest are similar in all aspects except for the condition being investigated. Moreover, phenotypic assessment of controls should be as comprehensive as the phenotypic assessment of cases. Ultimately, carefully crafted criteria, and thorough phenotypic assessments help reduce the impact of confounding variables.

Population stratification represents another potential source of confounding in candidate gene association studies utilizing a case-control design. The case-control design compares allele, genotype, or haplotype frequencies between the groups. Because these frequencies can be extremely disparate for different ancestries, it is important to control for ancestry to avoid spurious results/conclusions (e.g., concluding that there is an association between a phenotype/allele when in reality the association is fueled by ancestral differences in allelic frequencies). The risk for population stratification can be mitigated. Subgroup analysis represents one option, but it relies on self report to categorically measure race/ethnicity. An option that controls for population stratification statistically is the use of ancestral informative markers (AIM), which are polymorphisms in the DNA that allow one to calculate an admixture proportion for an individual. The application of these proportions are used for analysis rather than the traditionally used, though unreliable, method of self-reported race/ethnicity. In a recent study, only 30 AIMs were needed to estimate European admixture in a group of African American women (Ruiz-Narváez, Rosenberg, Wise, Reich, & Palmer, 2011). Although different AIMs may be needed to estimate other admixture proportions, this example demonstrates that population stratification can be successfully controlled through the analysis of genetic markers.

Another aspect of the candidate gene association study that should be considered is sample size requirements. Quanto (http://hydra.usc.edu/gxe/) is a freely downloadable computer program that can assist with sample size and/or power calculations for candidate gene association studies. User defined criteria can be manipulated according to the polymorphisms that have been selected for evaluation and according to study design specifications.

Candidate gene selection—Candidate gene selection is often based on biologic plausibility. This plausibility can be based on biological pathways implicated in the condition, biomarker data implicating a gene/gene product in the phenotype of interest, pharmacologic treatments for the condition that may indicate a target gene(s), or data from animal models (Hattersley & McCarthy, 2005). Bio-informatics databases, such as the Gene Ontology (http://www.geneontology.org/), may also aid in the identification of genes whose products may impact the phenotype of interest (The Gene Ontology, 1999–2011). Moreover, consideration should be given to number of genes on which to focus, ranging from a single gene to genes within a candidate biological pathway. Because more biologically global conclusions can be drawn, the study of a biologic pathway has the advantage of being more informative than the singular gene approach in most situations (Jorgenson, Ruczinski, Kessing, Smith, Shugart, & Alberg, 2009).

Polymorphism selection—Once selection of the candidate gene(s) is finalized, polymorphisms must be selected to evaluate candidate gene variability, and these are the genetic data used for analyses. The candidate gene association approach includes the

evaluation of single nucleotide polymorphisms (SNPs), repeat polymorphisms, insertion/deletion polymorphisms (INDEL), and copy number variants (CNV).

Resources for polymorphism selection: The SNP is the most common type of polymorphism and is a nucleotide (also known as a base) in the DNA where the nucleotide present (e.g., A, T, C, G) varies in the population (Genetics Home Reference, 2011). The scientific literature and a variety of online databases provide excellent resources for SNP identification and selection. A simple literature search combining the candidate gene(s) with the keyword "functional polymorphism" will help to identify SNPs known to alter the function of the candidate gene(s). Because functional polymorphisms modify the function of a gene regardless of phenotype, the literature search should not be limited to just the phenotype of interest. In addition to the literature, investigators also commonly use the Database of Single Nucleotide Polymorphisms (dbSNP) (http://www.ncbi.nlm.nih.gov/projects/SNP/) and the International HapMap Project (http://hapmap.ncbi.nlm.nih.gov/) to identify/select SNPs and tagging SNPs, respectively.

HapMap is accessed for the selection of tagging SNPs (tSNP), which represent the current gold standard for the evaluation of genetic variation in the candidate gene association study. The goal of HapMap is to develop a haplotype map of the human genome and to describe common patterns of genetic variation in humans (International HapMap Project, 2006). Essentially, HapMap is based on the premise that DNA is inherited in chunks/blocks (haploblock). Within these haploblocks, certain variants are inherited together. If the genotype of one variant within that block of DNA is known the genotype of a second variant within the same block can be determined since they are inherited together. Thus, HapMap assists the user in selecting SNPs that tag a certain haploblock of DNA (tagging SNPs or tSNPs). Ultimately, utilization of tSNPs allows one to fully evaluate the genetic variability of the candidate genes with the least number of SNPs (International HapMap Project, n.d.).

Repeat polymorphisms are characterized by repeating units of DNA bases. The number of times these DNA units repeat is variable in the population (Passarge, 2007). While repeat polymorphisms are less frequent in the genome than SNPs, they are often more informative as they usually have more alleles in the population than SNPs, which typically only have 2. The short tandem repeat (STR) is typically comprised of a repeating unit of two to four DNA bases (e.g., CAG CAG CAG) while the variable number tandem repeat (VNTR) is comprised of a larger repeating unit (Passarge), usually greater than 5 bases. For the evaluation of STRs and VNTRs, the literature continues to be the best source for identification and characterization.

An INDEL polymorphism occurs when a base(s) is added or subtracted from a place in the DNA. It is the presence or absence of the INDEL that is variable in the population (Nussbaum, McInnes, & Willard, 2007). Like SNPs, the dbSNP can be freely accessed to identify small-scale INDELs.

The CNV occurs when the number of copies of a particular genomic sequence/segment is variable in the population (NHGRI, n.d.). CNVs can be identified through scientific literature and online databases. The Database of Genomic Structural Variation (dbVar) (http://www.ncbi.nlm.nih.gov/dbvar) and The Copy Number Variation Project by the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/humgen/cnv/) are two online resources that may assist in CNV identification.

Genotype data collection technologies—Multiple options are available for SNP genotyping, including the polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) technique, real-time PCR allelic discrimination (e.g.

TaqMan®), multiplexing via mass spectrometry, and bead chip technology. Selection of the genotyping technique is guided by the number of samples and polymorphisms to be genotyped and available resources. PCR-RFLP, which is used to genotype SNPs based on differences in fragment lengths, is suitable when the number of SNPs and samples to be genotyped is relatively small. Real-time PCR allelic discrimination (http://www.appliedbiosystems.com; http://www.roche-applied-science.com), which genotypes SNPs based on allele-specific fluorescence intensity signals, is suitable for a medium number of SNPs and sample size. Because PCR-RFLP and real-time PCR allelic discrimination can only genotype one SNP at a time, the use of high throughput technologies have become the gold standard for SNP genotype collection when the number of SNPs to be evaluated approaches 24. The iPLEX® Gold-SNP Genotyping assay (http://www.sequenom.com), which genotypes SNPs based on differences in molecular mass, allows for the analysis of up to 36 SNPs per assay (Sequenom, 2010) in larger sample sizes. Not only can an investigator analyze multiple SNPs simultaneously, but time, assay to assay variability, and costs are reduced. The GoldenGate Genotyping Assay (http://www.illumina.com) is another high throughput bead based technology that can be utilized when the number of SNPs and samples to be analyzed is too large for other technologies.

There are several genotyping technologies also available for repeat polymorphisms, INDELs, and CNVs. PCR amplification followed by fragment sizing can be used for genotyping repeat polymorphisms. As with SNPs, real-time PCR allelic discrimination can be used to genotype small INDELs. Finally, TaqMan® Copy Number Assays (http://appliedbiosystems.com) or cytogenetic techniques (e.g., Fluorescence In Situ Hybridization) can be utilized for genotyping candidate CNVs.

Genome Wide Association Studies (GWAS)

Rationale for taking a GWAS approach—A GWAS gentoypes thousands to millions of polymorphisms across the genome for individuals who are phenotypically well-characterized (DiStefano & Taverna, 2011). If genetic variability is significantly different between cases and controls, those variations may be associated with susceptibility to or protection from the phenotype of interest and can provide direction as to which region of the genome these differences might be located. Ongoing efforts of the Human Genome Project and the International HapMap Project have made this approach possible through the generation of large databases that reference and map both sequence and variability.

The major advantage of a GWAS approach is that the biology of the phenotype of interest does not need to be completely understood prior to implementing this approach and the SNPs or genes of interest do not need to be defined *a priori*. Instead of selecting genes and polymorphisms *a priori*, polymorphisms that cover haploblocks across the entire genome are used for genotype data collection and non-parametric based analyses determine what genes/regions of the genome are relevant to the phenotype of interest (Hakonarson & Grant, 2011). The data derived from GWAS will provide direction regarding which areas of the genome warrant additional study.

There are several limitations to GWAS. The variant identified may not be what's accounting for the association, but is rather "tagging along" with the actual causal variant(s). This obstacle is also present for candidate gene association studies, particularly those that utilize a tSNP approach. Therefore, it may be necessary to follow up with more focused genotype data collection, including denser polymorphism evaluations and/or sequencing of that specific region of the genome to identify the exact allele accounting for the association (NHGRI, 2010). A major limitation for the GWAS approach, and perhaps a reason why many investigators are unable to pursue this approach, is the need for thousands of subjects

who are phenotypically well characterized and for which DNA is available. The need for large samples sizes for GWAS is due to the inherent issue of multiple testing that accompanies the evaluation of thousands to millions of different genetic variables. Additionally, the need for very large sample sizes, coupled with the cost of commercial genome-wide scanning techniques makes this approach very costly. GWAS approaches are also not optimal to assess rare polymorphisms as the data collection approaches for the GWAS are more focused on optimizing informativeness of the data (Ku, Loy, Pawitan, & Chia, 2010).

Subject and sample considerations—The cross-sectional case-control study design is the most frequently used approach for a GWAS. Study subjects should be selected based on a well-defined and heritable phenotype. Cases are defined as individuals who meet criteria for a phenotype of interest. Controls are individuals who have never met criteria for the phenotype and ideally have passed through the age or period of risk for the phenotype (Hakonarson & Grant, 2011). Like candidate gene associations studies, ancestry must be considered to avoid issues related to population substructure and this is why some investigators have conducted these types of studies with homogeneous populations (Psychiatric GWAS Consortium Coordinating Committee, 2009). Case and control groups should be matched on ancestry as much as possible to avoid false-positives. Despite this consideration, an advantage of GWAS is that whole genome data can provide adequate data to identify stratification and inflation of test statistics due to population substructure can be addressed (Hakonarson & Grant).

Obtaining a sufficiently large sample size is essential to ensure sufficient statistical power for a GWAS approach. Approximately 1,000 cases and a similar number of controls are required to detect 1–5 variants associated with a given trait. A larger sample is needed to uncover additional variants that may have diminishing contributions to the disease (Hakonarson & Grant, 2011).

Informed consent issues: While informed consent is of paramount importance with any research study, researchers who are considering a GWAS should be cognizant of issues related to conducting such as study and the National Institutes of Health (NIH) policy on data sharing for GWAS. In January 2008, the NIH adjusted its policy mandating the sharing of GWAS data obtained in NIH-funded or conducted studies. The details of this policy can be found at http://gwas.nih.gov/. Most NIH-funded GWAS are required to include language in the consent document that addresses public sharing of de-identified genotype and phenotype data. Researchers who are planning to study existing samples must ensure that the original consent signed by the subjects is consistent with conducting a GWAS.

Genotype data collection technologies—There are currently two commonly used vendors that provide technology for collection of GWAS data, Affymetrix and Illumina. The companies use different technological approaches, which are both widely used in the research community. The Affymetrix® Genome Wide SNP Array 6.0 features 1.8 million genetic markers, including 906,600 SNPs and more than 946,000 probes for the detection of CNVs. This platform also includes a high resolution reference map and a copy number polymorphism (CNP) algorithm (see http://www.affymetrix.com for additional information). The Illumina Omni Microarrays provide a multiple bead chip option which will soon include nearly 5 million markers per sample, including both common and rare variants identified by the 1000 Genomes project. Omni microarrays assess structural variation, including CNVs and copy neutral variants (inversions and translocations) which may also be significant contributors to disease (see http://www.illumina.com for additional information).

Resources of interest for GWAS: The Center for Inherited Disease Research (CIDR) at Johns Hopkins University (http://www.cidr.jhmi.edu/requirements/applications.html) is funded by NIH Institutes and provides genotyping and statistical genetic services to investigators who have received access after a competitive peer review process. Interested investigators are required to submit an application for projects supported by the NIH. In order to maximize access to resources, the application process to CIDR should ideally take place before or at the time of grant application, though this is not a requirement.

The repository for GWAS data is currently the Database of Genotypes and Phenotypes (dbGaP; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gap). This database was developed to archive the results of studies that have investigated the genotype-phenotype interaction and serves as a useful resource in reviewing the work that has already been completed and aids in planning future research. The dbGaP database provides the opportunity for *in silico* research. Researchers have the option of two levels of access (open and closed) to dbGaP: Open-access data are aggregate data that are publicly available while closed level access requires an application and approval process that includes de-identified subject specific data. The genotype data and their linked phenotype data are invaluable resources and researchers are encouraged to investigate this database as it pertains to their phenotypes of interest prior to designing a study.

GENE EXPRESSION STUDIES

Gene expression studies evaluate the activity of a gene using the level of messenger RNA (mRNA) from a gene(s) and determine if that level is associated with the phenotype of interest. DNA contains a code to generate mRNA through a process called transcription. The amount of mRNA produced from a gene, if at all, depends on many factors including tissue type, local cell environment, and point in the cell cycle.

A gene expression study is different from a polymorphism based study because an expression study evaluates mRNA levels that can change over time, uses less stable mRNA instead of DNA, and mRNA levels can be dramatically different based on what tissue is used for analysis, since gene expression is tissue-specific. Gene expression studies therefore should address whether multiple samples over time are needed for evaluation (similar to other types of biomarkers that change over time), RNA stabilization, and what cell/tissue type is most appropriate to evaluate for the phenotype of interest. For these reasons, many stored samples may not be appropriate for this approach.

Candidate Gene Expression Studies

Rationale for taking a candidate gene expression approach—Candidate gene expression studies investigate mRNA levels for a specific gene(s) to determine if it is associated with a phenotype of interest. Similar to a candidate gene association approach, this is a hypothesis-driven approach where the investigator *a priori* selects the candidate gene(s) to be evaluated. This approach is only appropriate if the investigator has ample justification for investigating a specific gene(s).

Subject and sample considerations—Gene expression studies often involve relative comparisons of mRNA levels between two groups (these groups can be different types of tissues, groups that vary for a particular exposure or groups that vary by the presence or absence of a phenotype of interest). Clearly defined inclusion/exclusion criteria are necessary, due to the relative comparison nature of this approach.

RNA stabilization: Stabilization of RNA is essential to obtain accurate gene expression profiles of biological samples. Immediately after sample collection, RNA degradation and

other transcriptional changes begin to occur. These alterations may result in false up or down regulation of gene expression levels. RNA stabilization preserves a representative gene expression profile for later analysis (e.g. quantitative RT-PCR and microarray analysis). RNA stabilization methods vary based on the type of biological sample. Five of the most common RNA stabilization methods are: (a) PAXgene Blood RNA System (http://www.preanalytix.com), which utilizes a single tube (pre-filled with RNA stabilization reagent) for blood collection, RNA stabilization, sample transport and storage, and purification of total RNA (PreAnalytiX, 2010); (b) LeukoLOCK System (http:www.lifetechnologies.com), which filters and isolates leukocytes from whole blood. RNAlater solution is then used to stabilize the RNA of the leukocytes. A notable advantage to the LeukoLOCK System is the ability to remove a large proportion of reticulocyte-derived globin mRNA. Depletion of the globin mRNA allows for the detection of thousands of additional genes on microarray (Life Technologies Corporation LeukoLOCK, 2010); (c) RNAlater (http://www.ambion.com; http:www.qiagen.com) stabilizes RNA in a variety of fresh samples including animal tissue, tissue culture cells, leukocytes, yeast, and bacteria. After collection, the sample is submerged in the RNAlater stabilization solution. This solution permeates and stabilizes the sample eliminating the need for immediate processing or freezing of samples (Life Technologies Corporation RNAlater, 2010; Qiagen, 2006); (d) Oragene RNA for Expression Analysis Self Collection Kit (http://www.dnagenotek.com): allows for the non-invasive collection of RNA from saliva. Donors are instructed to expectorate into a vial, cap the container, and shake vigorously to release a stabilization solution from the cap. Oragene RNA samples can remain stable for months at room temperature (DNA Genotek, 2011); and (e) Snap Freeze quick freezes solid tissues with liquid nitrogen and dry ice can be used to preserve RNA; however, disruptions during freezing and thawing can lead to RNA degradation. Due to potential RNA degradation and difficulty of obtaining and working with liquid nitrogen and dry ice, RNAlater described above may be a more viable option for solid tissue RNA stabilization.

Candidate gene selection—Candidate gene selection must be justified and rationale for selection is similar to selection of candidate genes in the candidate gene association section. The same bio-informatics databases mentioned in that section are also applicable to aiding in the selection of candidate genes for an expression study and as with a candidate gene association study, a candidate gene expression study should consider focusing on a group of genes in a biological pathway versus the value of focusing on a single gene.

Expression data collection technologies—Selection of a data collection technology for a candidate gene expression study should take into account the number of genes/loci and the number of samples to be evaluated. The most frequently used technologies for a candidate gene approach include Northern blotting, quantitative real-time PCR (qRT-PCR), and multiplex platforms that support 3–36 genes/loci per reaction.

Northern blotting requires electrophoresis of RNA, transfer to a membrane and hybridizing the membrane with a probe specific for detection of the mRNA of interest. The advantages of blotting are that most laboratories will have the equipment to conduct this type of data collection and assessment of RNA size is possible. Disadvantages of blotting are that RNA degradation is common, it requires more RNA as a template for the experiment compared to other methods, it is laborious, and is not optimal for quantification of mRNA levels.

Currently, one of the most popular techniques for assessing the level of mRNA for a gene/locus is qRT-PCR. qRT-PCR requires conversion of RNA to a more stable template called cDNA (complementary DNA), PCR amplification and probe hybridization for the gene/locus of interest. The probe is fluorescently labeled and liberation of this fluorescent label is quantified, reflecting the amount of starting mRNA template in the sample. One crucial step

in conducting qRT-PCR is normalization of the data generated. Normalization of the data allows for sample to sample comparisons that have been corrected for noise such as what's introduced when sample dispensing between samples isn't uniform. This is often done using qRT-PCR data collected simultaneously for an endogenous control, which usually represents a stably expressed gene (often referred to as a "housekeeping gene") and allows for normalization of data across samples (Guenin et al., 2009). Thought needs to be given to selection of an appropriate endogenous control given that different tissues will have different stably expressed genes (Guenin et al.). If in doubt, endogenous control panels are available for assessment prior to conducting qRT-PCR. Advantages of qRT-PCR include high sensitivity and reduced RNA template requirements, high throughput capabilities, quantification of starting mRNA template is possible with use of proper exogenous reference controls, and for many genes/loci/pathways off the shelf optimized assays are available (http://www.appliedbiosystems.com; http;//www.roche-applied-science.com).

Multiplex gene expression assays are available when the number of genes/loci to be evaluated is in the range of ~3–36. One example is the QuantiGene® Plex 2.0 assay (for more information see http://www.panomics.com/index.php?id=product_6) that uses Luminex technology to collect the data and the assay can be customized.

Global (Genome Wide) Gene Expression Studies

Rationale for taking a global gene expression approach—Whole genome expression (also known as global gene expression or gene expression profiling) offers a comprehensive view of gene activity within a biological sample by examining mRNA levels for all known genes across the genome. In this way, whole genome expression provides functional information regarding "when and where a protein is expressed, when it is degraded, and with which other proteins it may interact" (Altman & Raychaudhuri, 2001, p. 340). Due to the dynamic nature of expression, gene expression profiles are often relatively compared under multiple conditions (such as comparing different tissue types, comparing normal versus abnormal tissues, comparing tissues before and after an exposure) or over a period of time (Altman & Raychaudhuri, 2001; Arcellana-Panlilio & Robbins, 2002). The use of global gene expression profiling is extremely advantageous when little to nothing is known about the genes influencing a condition, a similar advantage held by the GWAS approach. Thus, whole genome expression can identify novel candidate hypotheses through a non-parametric analysis of genome wide expression data.

Subject and sample considerations

Sample selection: Although this is an approach similar to GWAS, with evaluation of thousands of genes in a nonparametric manner, sample size requirements for global gene expression are usually smaller, requiring approximately 10 subjects per variable. Matching of subjects for key variables known to influence the phenotype under investigation can reduce the number of variables that need to be accounted for in the analyses. A sample size calculator for global gene expression experiments can be found at http://bioinformatics.mdanderson.org/MicroarraySampleSize/. Additionally, as with candidate gene expression studies, mRNA stabilization of the collected samples is crucial.

Gene expression data collection technologies

Microarrays: Microarrays are used to examine the expression profile of a single sample (often referred to as single dye array) or to compare expression levels between two different samples/conditions (often referred to as two dye array). The microarray itself is a solid surface covered with an "ordered arrangement of unique nucleic acid fragments derived from individual genes" (Arcellana-Panilio & Robbins, 2002, p. G397). Fluorescently labeled template hybridizes to these nucleic acid fragments (referred to as probes) on the solid

surface through complementary pairing. The intensity of the florescence at each spot on the microarray corresponds to the amount of sample binding to a particular nucleic acid fragment and thus, the gene expression level. If the microarray reveals any interesting findings, q-RT-PCR should be carried out for validation purposes. For a visual representation of microarray methodology visit this web address: http://www.bio.davidson.edu/courses/genomics/chip/chip.html.

Microarrays have revolutionized gene expression analyses, as this technology is able to simultaneously survey thousands of genes in a short period of time. However, the ability to detect novel genes is limited to the hybridization probes represented on the microarray. Off-the-shelf probe sets that contain reference sequences can be used, or custom probe sets are designed based on specific genes of interest or pathways. Additionally, microarrays require specialized lab equipment and are very useful when analyzing a small sample size but become costly as sample size increases. Two popular microarrays platforms include Affymetrix's GeneChip and Illumina's BeadChip.

Affymetrix's GeneChip platform (for more information see http://www.affymetrix.com) utilizes traditional solid support microarray technology. Affymetrix's latest product, the GeneChip Human Gene 1.0 ST Array, is able to interrogate 28,869 genes and covers over 700,000 distinct probes. A greater number of samples can be processed simultaneously (with this same probe set) using the Human Gene 1.1 Array Strip (4 samples/strip) and the Human Gene 1.1 Array Plate (16, 24, or 96 samples/plate). Affymetrix also provides whole transcript expression analysis technology for mice and rats.

Instead of using a solid support platform, the Illumina BeadChip platform (for more information see http://www.illumina.com) employs silica beads (each covered with thousands of copies of a specific oligonucleotide) self-assembled in microwells of fiber optic bundles or planar silica slides. Illumina's most recent whole genome expression array, the HumanHT-12 v4 BeadChip, provides high throughput processing of twelve samples and covers over 47,000 probes. Illumina also offers whole genome expression BeadChip technology for mice and rats.

Normalization of gene expression data is also important with microarray data collection. Unlike qRT-PCR where an appropriate endogenous control needs to be selected and included in the data collection, microarrays already include a range of endogenous controls for which data is simultaneously collected and from which the investigator can select to use for normalization of the data.

Sequence based technologies that utilize next-generation sequencing (NGS; high throughput sequencing) are also available for collection of genome-wide gene expression data. An example of such a technology is the RNA-Seq method (for more information see http://www.illumina.com). This method requires conversion to cDNA, ligation of the cDNA fragments, creation of a library, sequencing of the template, and collection of frequency data for a transcript. An advantage of this approach over microarrays is that it does not require primers or probes therefore novel transcripts that would not be detectable with a microarray can be identified.

Serial analysis of gene expression: Serial analysis of gene expression (SAGE) provides comprehensive quantitative gene expression data. SAGE technology is based on three main principles: (1) a short sequence tag (9–17 bases) contains sufficient information to distinctively identify a transcript, (2) sequence tags can be linked together to form one long molecule that can be cloned and sequenced to allow efficient analysis of transcripts, and (3) the number of times a particular tag is observed corresponds to the expression level of the

transcript (Sagenet, 2005; Velculescu, Zhang, Vogelstein, & Kinzler, 1995). One of the main advantages of SAGE, similar to RNA-Seq, is the ability to detect novel genes as it does not require prior sequence information or hybridization probes for each transcript like microarrays (Velculescu et. al, 1997). Another advantage of SAGE is that is utilizes common laboratory equipment and techniques. Any laboratory that performs PCR and manual sequencing could also execute SAGE. Nonetheless, due to cloning and sequencing, SAGE can be expensive, time consuming, and labor intensive.

EPIGENETIC STUDIES

An epigenetic mechanism is a biochemical alteration to the DNA molecule that does not change the sequence of the DNA but does influence gene expression. Epigenetics is often defined as the "study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (Russo, Martienssen, & Riggs, 1996, p. 1).

The epigenetic/epigenomic approach shares many advantages and disadvantages with DNA polymorphism based approaches and gene expression based approaches. Like DNA polymorphism based approaches, the epigenetic/epigenomic approach uses DNA as its template for data collection. Since both DNA sequence and its chemical modifications are stable, stored samples are more likely to be appropriate for this approach than gene expression approaches. Similar to a gene expression based approach; epigenetic/epigenomic alterations can change over time and can differ dramatically between cell/tissue types. Although template stability is not an issue, the investigator should give great consideration to whether multiple samples over time are needed for evaluation and what cell/tissue type is most appropriate to evaluate for the given phenotype of interest. For these reasons, similar to gene expression studies, many stored samples may not be appropriate for this approach.

Chromatin remodeling, non-coding RNAs, histone modifications, and DNA methylation are all epigenetic/epigenomic alterations that impact gene expression. Chromatin remodeling is an enzymatic process that results in altered chromatin and nucleosome composition. This transformed structure provides regulatory proteins access to the DNA molecule. Non-coding RNAs are not translated into protein but have considerable involvement in gene expression through interactions with DNA/mRNA. While chromatin remodeling and non-coding RNAs are important to gene regulation, this paper will focus primarily on the commonly examined epigenetic mechanisms for which the most technology for data collection is available: histone modifications and methylation.

Rationale for Taking an Epigenetic/Epigenomic Approach

The decision to take an epigenetic (candidate gene) or an epigenomic (genome wide) approach is based upon wanting to evaluate the mechanism for gene regulation. There are many environmental factors that impact the severity and frequency of epigenetic/epigenomic alterations and subsequent gene expression; therefore, this approach is often used to examine multifactorial diseases that have an environmental component associated with it. Epigenetic approaches to examine transcriptional regulation have contributed to a more comprehensive understanding of complex conditions that demonstrate aberrant gene expression, including: cancer (Wilop et al., 2011), mental health (Read, Bentall, & Fosse, 2009), and cardiovascular disorders (Ordovás & Smith, 2010). Furthermore, the investigation of diseases for which DNA mutations have not been revealed may benefit from an epigenetic approach.

Subject and Sample Considerations

The epigenome is subject to frequent alterations; therefore, longitudinal sample collection is recommended if evaluating time-sensitive trends. Subject size recommendations for an epigenetic study follow similar guidelines to a gene expression study, and vary on whether the investigator will examine the entire genome (hypothesis generating/larger sample size) or a candidate gene profile (hypothesis driven/smaller sample size). Like the other approaches described, an epigenetic study does not require that subjects be related. The advantages and disadvantages of conducting a genome-wide versus candidate gene epigenetic study are similar to those described in previous sections.

The epigenome is largely determined by cell type, and this is especially true for methylation patterns; therefore, tissue source is extremely important to consider for this type of approach. For example, the methylation profile of a skin cell is very different than the methylation profile of a liver cell, since different genes are expressed in each cell type, and methylation is a driving force behind tissue specific gene expression. Similar to a gene expression study, an epigenetic design requires the samples for epigenetic analyses be from a tissue that appropriately addresses the phenotype of interest. Tissue specific sample collection will capture epigenetic patterns that impact gene expression which are potentially contributing to the disease. Unlike a gene expression study which examines RNA, this design requires DNA, which is advantageous for the investigator who has access to previously collected samples, assuming they were collected from an appropriate tissue for the phenotype under investigation.

Epigenetic and Epigenomic Data Collection Technologies

This section will focus on the two epigenetic mechanisms most frequently studied: (a) histone modification and (b) methylation. Post-translational histone modifications include alteration of the histone tail through biochemical changes that ultimately impact gene activity. Genome-wide histone modifications can be captured with chromatin immunoprecipitation technology (ChIP), and quantified with a microarray (ChIP-chip). Methylation refers to the addition of a methyl group to a cytosine, often at CpG islands, which are regions of the genome that are rich in CG base sequences. Hypermethylation of a gene typically leads to gene suppression, while hypomethylation results in gene expression. Genome-wide methylation intensities can also be measured with affinity-based immunoprecipitation (MeDIP), and quantified with a microarray (MeDIP-chip, Infinium platform). Methylation of candidate genes can also be measured with restriction enzymes that recognize only demethylated CpG regions (HELP assay), or pyrosequencing. Next Generation Sequencing approaches are also becoming increasingly popular, more cost effective, and provide global sequencing for histone modification (ChIP-seq) and methylation (MeDIP-seq), often integrating these with other epigenetic mechanisms. This section will describe each method and provide the reader with technologies and recommendations to aide in the design and implementation of an epigenetic study.

Histone modification analysis

Histone modification signals can be captured with chromatin immunoprecipitation (ChIP), which provides modification position approximation on the genome (Collas, 2010). The ChIP-chip technique combines this ChIP technology with a microarray (chip) to quantify the sum of binding sites on the genome (Aparacio, Geisberg, & Struhl, 2004). The ChIP-seq technique (see Next Generation Sequencing) has become a popular technique compared to ChIP-chip. Unlike ChIP-seq, ChIP-chip requires more amplification, multiplexing is not possible (Park, 2009), and the results have a lower resolution that are limited to the coverage provided by the selected microarray (Evertts, Zee, & Garcia, 2010). Nimble Gen offers a whole-genome ChIP-chip tiling array that allows the investigator to choose between

ordering the entire genome set or individual arrays within a set (http://www.nimblegen.com/products/chip/wgt/index.html). Single gene ChIP technologies are available that target antibodies against specific histone modifications. Mass spectrometry also allows the measurement of mass-to-charge ratio of peptides (Evertts et al.) and allows for changes in modification to be quantified during chromatin assembly (Deal & Henikoff, 2010).

When performing any microarray experiment, it is important to address concerns that may compromise the integrity of the experiment, including: image acquisition, background subtraction, standard normalization and the need to control for biases from dye (Buck & Lieb, 2004). Additionally, the reproducibility of the histone-modification results depends on the quality and specificity of antibodies used. Antibodies may exhibit appropriate specificity, but are ineffective when subjected to ChIP reagents (Egelhoffer et al., 2010). The Center for Biomedical Informatics at Harvard Medical School has developed an online repository that allows investigators to search for antibodies subjected to validation tests (http://compbio.med.harvard.edu/antibodies/about). It is important to note that this validation data should be used as a guide and investigators are encouraged to validate their own findings.

Bisulfite-conversion based methylation analyses

Bisulfite-conversion of unmethylated cytosines to uracils remains the gold-standard to evaluate methylation (Huang, Huang, & Feng, 2010). Bisulfite-conversion based microarrays use probes that hybridize targets to methylated and unmethylated regions, and release a fluorescent intensity that denotes methylation status (Huang et al.). Recent research indicates that tissue-specific methylation occurs in CpG island shores rather than previously targeted CpG islands (Irizarry et al., 2009); therefore, CpG islands alone are not sufficient to reveal differentially methylated regions and methylome evaluation should also include CpG shores (Gupta, Nagarajan, & Wajapeyee, 2010). Like other non-sequencing-based methods, the results of this platform are "susceptible to certain polymorphisms that were not known or considered at the time the array was designed" (Rakyan, Down, & Balding, 2011, p. 532). Illumina offers the Infinium HumanMethylation450K which provides a whole-genome analysis of methylation intensities of more than 450,000 sites, including CpG islands, shores and other CpG sites outside of islands (for more information see http://www.illumina.com/products/methylation_450_beadchip_kits.ilmn). Candidate gene methylation assessment can be accomplished through technologies such as the EpiTYPER (for more information see http://www.sequenom.com) that uses bisulfite converted DNA as a template for PCR and after modification and cleavage of the PCR product, mass spectrometry is performed to quantify methylated and non-methylated DNA.

Bisulphite-based sequencing (BS-seq) uses bisulphite converted DNA as a template, PCR amplification occurs, and sequencing of the resulting fragments provide a global view of methylation with minimal bias toward CpG dense regions. This approach provides the highest level of coverage and resolution, but is not capable of distinguishing between methylated and hydroxymethylated cytosine bases. BS-seq can be used for both a genome-wide or candidate gene approach. Pyrosequencing examines the methylation intensity of specific sites or genes of interest. Illumina offers a single-site resolution methylation assay that uses bisulfite conversion and pyrosequencing to produce high resolution results (http://www.illumina.com/technology/veracode_methylation_assay.ilmn).

Affinity-based methylation analyses

Genome-wide affinity-based microarrays use enzyme recognition sites within CpG sites that enrich the methylated fraction of the genome. The MeDIP-chip technique (Methylated DNA

Immunprecipitation-chromatin Immunoprecipitation) immunoprecipitates the methylated portion of genomic DNA with an antibody, and is followed by quantification of methylation with a microarray. This technique yields a restricted resolution that is limited by the type of array used. MeDIP-chip should be validated with quantitative PCR, though referencing is not required since bisulfite conversion does not occur. ArrayStar offers MeDIP-chip services that include quality assessments for both methods (http://www.arraystar.com/Microarray/service_main.asp?id=181).

Restriction endonuclease-based methylation analysis

Restriction endonucleases have been adapted to discriminate methylated from unmethylated regions in the DNA (Edwards et al., 2010). This approach uses restriction enzymes that recognize only unmethylated sites, and are therefore unable to cut methylated portions of DNA. This method, combined with high throughput sequencing is limited by the availability of restriction enzyme sites in the target DNA (Gupta et al., 2010). Additionally, this technique requires large amounts of DNA (Biotage, 2007). Advantages for this approach include: a simplified data analysis, straightforward protocol, and it does not require bisulfite-conversion. The use of restriction enzymes to analyze methylation can be used for either candidate-gene or genome-wide studies (Gupta et al.) and has been used as a method of methylation mapping analysis (Edwards et al., 2010).

Data Quality assessments are important to incorporate into an epigenetic study. Quantile and LOESS normalization is recommended, which assumes a similar total strength (source). Additionally, bisulfite-based experiments, especially pyrosequencing since PCR is highly variable, should include verification in independent samples to distinguish methylation from incomplete bisulfite conversion (Laird, 2010). Incomplete conversion of methylated cytosines remains a major weakness of bisulfite-conversion based analysis techniques. Fully methylated and fully unmethylated controls should be provided by commercial vendors which allow the investigator to evaluate bisulphite-conversion efficiency.

Next generation sequencing (NGS) for histone modification analysis

DNA sequencing from epigenetic events may provide a first step toward quantification of epigenetic mechanisms. Similar to ChIP-chip, ChIP-seq uses antibodies to enrich for histone modifications, but is instead followed by high-throughput sequencing that measures gene expression levels (Everts et al., 2010). This technique determines the genome-wide patterns of modified chromatin, including: histone methylation, acetylation status and binding regions for proteins (Werner, 2010). Unlike ChIP-chip, ChIP-Seq offers higher resolution with fewer artifacts, greater coverage, and requires less DNA. Illumina offers a ChIP-seq assay that provides a wide range of binding sites with varying strength (http://www.illumina.com/technology/chip_seq_assay.ilmn).

Next generation sequencing (NGS) for methylation analysis

MeDIP-seq (Methylated DNA Immunoprecipitation-Sequencing) is a high throughput sequencing technique of methylated DNA fragments that is aligned to a referenced genome. This technique is comparatively easier to analyze and interpret (Gupta et al., 2010); however, this method is best used to study hypermethylation of CpG-rich areas, since methylated CpG-rich sequences are more efficiently enriched than methylated CpG-poor sequences (Bibkova & Fan, 2009).

CONCLUSIONS

Nurse scientists should give much thought to how a genetic or genomic study could positively impact and move forward their program of research. When designing a genetic or

genomic research study it is paramount that one decides if they will take a polymorphism based, gene expression based or epigenetic based approach and then within the context of that study whether they will take a genetic or a genomic approach. This paper, while not providing an exhaustive review of available technologies, demonstrates the variety of technologies available for commonly used approaches, each with advantages and disadvantages. Availability of databases housing information to facilitate study design, data collection, interpretation of findings, and dissemination of data have greatly improved over the past decade. Investigators are encouraged to visit and utilize *in silico* resources when designing a research study to ensure they are conducting novel investigations and using up to date information.

Acknowledgments

The authors would like to acknowledge support available through a National Institutes of Health, National Institute of Nursing Research award "Targeted Research and Academic Training Program for Nurses in Genomics" (T32 NR009759).

References

Altman RB, Raychaudhuri S. Whole-genome expression analysis: challenges beyond clustering. Current Opinion in Structural Biology. 2001; 11(3):340–347.10.1016/S0959-440X(00)00212-8 [PubMed: 11406385]

Aparicio, O.; Geisberg, J.; Struhl, K. Current Protocols in Cell Biology. Vol. Chapter 17. Los Angeles, CA, USA: John Wiley & Sons, Inc; 2004. Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo.

Arcellana-Panlilio M, Robbins SM. Global gene expression profiling using DNA microarrays. American Journal of Physiology – Gastrointestinal and Liver Physiology. 2002; 282(3):G397–G402.10.1152/ajpgi.00519.2001 [PubMed: 11841989]

Bibkova M, Fan J. Genome-wide DNA methylation profiling. WIRE Systems Biology and Medicine. 2009; 2(2):210–223.10.1002/wsbm.35

Biotage. CpG methylation analysis by pyrosequencing: benchmarks and application. 2007 Mar. Retrieved July 11, 2011, from http://www.pyrosequencing.com/graphics/7424.pdf

Buck MJ, Lieb JD. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. Genomics. 2004; 83(3):349–360.10.1016/j.ygeno.2003.11.004 [PubMed: 14986705]

Campbell, AM. Molecular movies: DNA microarray methodology. 2001. Retrieved July 13 2011 , from http://www.bio.davidson.edu/courses/genomics/chip/chip.html

Collas P. The current state of chromatin immunoprecipitation. Molecular Biotechnology. 2010; 45(1):87–100.10.1007/s12033-009-9239-8 [PubMed: 20077036]

Corvin A, Craddock N, Sullivan PF. Genome-wide association studies: A primer. Psychological Medicine. 2010; 40(7):1063–1077.10.1017/S0033291709991723 [PubMed: 19895722]

Deal R, Henikoff S. Capturing the dynamic epigenome. Genome Biology. 2010; 11(10):218.10.1186/gb-2010-11-10-218 [PubMed: 20959022]

DiStefana, JK.; Taverna, DM. Technical issues and experimental design of gene association studies. In: DiStefano, Joanna K., editor. Disease Gene Identification: Methods and Protocols, Methods in Molecular Biology. 2011. p. 3-16.

DNA Genotek. Oragene RNA. 2011. Retrieved July 13, 2001, from http://www.dnagenotek.com/DNA_Genotek_Product_RNA_Overview.html

Edwards JR, O'Donnell AH, Rollins RA, Peckham HE, Lee C, Milekic MH, Bestor TH. Chromatin and sequence features that define the fine and gross structure of genomic methylation patterns. Genome Research. 2010; 20(7):972–980.10.1101/gr.101535.109 [PubMed: 20488932]

Egelhoffer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Lieb JD. An assessment of histone-modification antibody quality. Nature Structural & Molecular Biology. 2010; 18(1):91–93.10.1038/nsmb.1972

Everts A, Zee B, Garcia B. Modern approaches for investigating epigenetic signaling pathways. Journal of Applied Physiology. 2010; 109(3):927–933.10.1152/japplphysiol.00007.2010 [PubMed: 20110548]

Genetics Home Reference. What are single nucleotide polymorphisms (SNPs)?. 2011. Retrieved July 19, 2011, from http://ghr.nlm.nih.gov/handbook/genomicresearch/snp

Grant S, Hakonarson H. Microarray technology and applications in the arena of genome-wide association. Clinical Chemistry. 2008; 54(7):1116–1124.10.1373/clinchem.2008.105395 [PubMed: 18499899]

Green ED, Guyer MS. National Human Genome Research Institute. Charting a course for genomic medicine from base pairs to bedside. Nature. 2011; 470(7333):204–213.10.1038/nature09764 [PubMed: 21307933]

Guenin S, Mauriat M, Pelloux J, Van Wuytswinkel O, Bellini C, Gutierrez L. Normalization of qRT-PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. Journal of Experimental Botany. 2009; 60(2):487–493.10.1093/jxb/ern305 [PubMed: 19264760]

Gupta R, Nagarajan A, Wajapeyee N. Advances in genome-wide DNA methylation analysis. Biotechniques. 2010; 49(4):iii–xi.10.2144/000113493 [PubMed: 20964631]

Hakonarson H, Grant S. Planning a genome-wide association study: Points to consider. Annals of Medicine. 2011; 43(6):451–460.10.3109/07853890.2011.573803 [PubMed: 21595511]

Hattersly AT, McCarthy MI. Genetic Epidemiology 5: What makes a good genetic association study? The Lancet. 2005; 366(9493):1315–1323.10.1016/S0140-6736(05)67531-9

Huang Y, Huang T, Wang L. Profiling DNA methylomes from microarray to genome-scale sequencing. Technology in Cancer Research and Treatment. 2010; 9(2):139–147. Retrieved from PubMed database. [PubMed: 20218736]

International HapMap Project. About the International HapMap Project. 2006. Retrieved July 18, 2011, from http://hapmap.ncbi.nlm.nih.gov/abouthapmap.html

International HapMap Project. What is the HapMap?. n.d. Retrieved July 18, 2011, from http://hapmap.ncbi.nlm.nih.gov/whatishapmap.html.en

Irizarry RA, Ladd-Acosta B, Wen Z, Wu C, Montano P, Onyango H, Feinberg AP. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nature Genetics. 2009; 41(2):178–186.10.1038/ng.298 [PubMed: 19151715]

Jorgensen TJ, Ruczinski I, Kessing B, Smith MW, Shugart YY, Alberg AJ. Hypothesis-drive candidate gene association studies: Practical design and analytical considerations. American Journal of Epidemiology. 2009; 170(8):986–993.10.1093/aje/kwp242 [PubMed: 19762372]

Ku CK, Loy EY, Pawitan Y, Chia KS. The pursuit of genome-wide association studies: Where are we now? Journal of Human Genetics. 2010; 55(4):195–206.10.1038/jhg.2010.19 [PubMed: 20300123]

Laird P. Principles and challenges of genome-wide DNA methylation analysis. Nature Reviews Genetics. 2010; 11:191–203.10.1038/nrg2732

Lander ES. Initial impact of the sequencing of the human genome. Nature. 2011; 470(7333):187–197.10.1038/nature09792 [PubMed: 21307931]

Life Technologies Corporation. LeukoLOCK total RNA isolation system. 2010. Retrieved July 13, 2011, from http://www.ambion.com/techlib/prot/fm_1923.pdf

Life Technologies Corporation. RNAlater tissue collection: RNA stabilization solution. 2010. Retrieved July 13, 2011, from http://www.ambion.com/techlib/prot/bp_7020.pdf

National Center for Biotechnology Information. Microarrays: Chipping away at the mysteries of science and medicine. 2007. Retrieved June 6, 2011, from http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html

National Human Genome Research Institute [NHGRI]. Talking Glossary of Genetic Terms: Copy Number Variation (CNV). n.d. Retrieved July 19, 2011, from http://www.genome.gov/glossary/index.cfm?id=40

National Human Genome Research Institute [NHGRI]. Talking Glossary of Genetic Terms: Polymorphism. n.d. Retrieved July 26, 2011, from http://www.genome.gov/glossary/index.cfm?id=160

Nussbaum, RL.; McInnes, RR.; Huntington, FW. Thompson & Thompson genetics in medicine. 7. Philadelphia, PA: Saunders Elsevier; 2007.

Ordovás JM, Smith C. Epigenetics and cardiovascular disease. Nature Reviews Cardiology. 2010; 7(9):510–519.10.1038/nrcardio.2010.104

Park P. ChIP-seq: advantages and challenges of a maturing technology. Nature Reviews. 2009; 10(10): 669–680.10.1038/nrg2641

Passarge, E. Color Atlas of Genetics. New York: Thieme; 2007.

PreAnalytiX. PAXgene blood RNA: The better the source, the more to explore. 2010. Retrieved July 13, 2011, from http://www.qiagen.com/literature/render.aspx?id=200337

Psychiatric GWAS Consortium Coordinating Committee. Genomewide Association Studies: History, rationale, and prospects for psychiatric disorders. American Journal of Psychiatry. 2009; 166(5): 540–556.10.1176/appi.ajp.2008.08091354 [PubMed: 19339359]

Qiagen. RNAlater handbook. 2006. Retrieved July 13, 2011, from www.qiagen.com/literature/render.aspx?id=403

Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nature Reviews Genetics. 2011; 12(8):529–541.10.1038/nrg3000

Read J, Bentall R, Fosse R. Time to abandon the bio-bio-bio model of psychosis: Exploring the epigenetic and psychological mechanisms by which adverse life events lead to psychotic symptoms. Epidemiologia e Psichiatria Sociale. 2009; 18(4):299–310.10.1017/S1121189X00000257 [PubMed: 20170043]

Ruiz-Naváez EA, Rosenberg L, Wise LA, Reich D, Palmer JR. Validation of a small set of ancestral informative markers for control of population admixture in African Americans. American Journal of Epidemiology. 2011; 173(5):587–592.10.1093/aje/kwq401 [PubMed: 21262910]

Russo, VA.; Martienssen, RA.; Riggs, AD. Epigenetic Mechanisms of Gene Regulation. Vol. 32. Plainview, NY: Cold Spring Harbor Laboratory Press; 1996.

Sagenet. Description of SAGE. 2003–2005. Retrieved June 6, 2011, from http://www.sagenet.org/findings/index.html

Sequenom. MassARRAY® iPLEX® gold-SNP genotyping: From target discovery to HTP validating (version 2) [Brochure]. 2010.

The Gene Ontology. An Introduction to the Gene Ontology. 1999–2011. Retrieved July 19, 2011, from http://www.geneontology.org/GO.doc.shtml

The University of Texas MD Anderson Cancer Center: Department of Bioinformatics and Computational Biology. Sample size for microarray experiments. 2003–2010. Retrieved July 13, 2011, from http://bioinformatics.mdanderson.org/MicroarraySampleSize/

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. Serial analysis of gene expression. Science. 1995; 270(5235):484–487.10.1017/S1121189X00000257 [PubMed: 7570003]

Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE, Kinzler KW. Characterization of the yeast transcriptome. Cell. 1997; 88:243–251.10.1016/S0092-8674(00)81845-0 [PubMed: 9008165]

Werner T. Next generation sequencing in functional genomics. Briefings in Bioinformatics. 2010; 2(5):499–511.10.1093/bib/bbq018 [PubMed: 20501549]

Wilop S, Fernandez AF, Jost E, Herman JG, Brummendorf TH, Esteller M, Galm O. Array-based DNA methylation profiling in acute myeloid leukaemia. British Journal of Haematology. 2011; 155(1):65–72.10.1111/j.1365-2141.2011.08801.x [PubMed: 21790528]

Table 1

Online Genome Databases and Resources

Name and Address	Description
Database of Short Genetic Variations (aka SNP database) http://www.ncbi.nlm.nih.gov/snp/?term=	This database houses documented SNPs, microsatellites, and small-scale INDELs. It provides population specific allele frequencies; genotype data, genome location, and information on function (e.g., change in an amino acid).
International HapMap Project http://hapmap.ncbi.nlm.nih.gov/	This database is used to identify and select tagging SNPs. User defined criteria under the configure tab include population selection, R^2 cutoff values, and mean allele frequency cutoff. SNPs identified in the literature or dbSNP can also be included in the tagger SNP configuration.
Database of Genomic Structural Variation http://www.ncbi.nlm.nih.gov/dbvar	This database houses information on documented structural variants, including CNVs. User defined limits include criteria such as study design, method type (e.g., SNP genotyping, FISH), project ID, and variant type
Copy Number Variation (CNV) Project http://www.sanger.ac.uk/humgen/cnv/	This database provides CNV data from two projects (Global CNV assessment; High-resolution CNV discovery)
Genetics Home Reference http://ghr.nlm.nih.gov/	This website by the National Library of Medicine contains information concerning genetic conditions, genes, and chromosomes.
Talking Glossary of Genetic Terms http://www.genome.gov/glossary/index.cfm	This glossary provides definitions, illustrations, and animations of commonly used genetic/genomic terms.
The Gene Ontology Project http://www.geneontology.org/	This database can be used to identify genes whose products may impact a phenotype of interest. The domains covered include cellular component, molecular function, and biological process.
Catalog of Published Genome-Wide Association Studies http://www.genome.gov/gwastudies/	Database containing all published GWA studies attempting to genotype at least 100,000 SNPs in the initial stage
Genome-Wide Association Studies Data Repository http://was.nih.gov/	Website for the NIH Genome Wide Association Study Portal
The Genes, Environment, and Health Initiative http://www.genesandenvironment.nih.gov	Website for Genes, Environment and Health Initiative (GEI)
Database of Genotypes and Phenotypes http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gap	Database containing results of studies investigating genotype-phenotype interaction. Currently houses NIH GWAS repository.
Center for Inherited Disease Research http://www.cidr.jhmi.edu	Provides genotyping and statistical genetic services to investigators approved for access through competitive peer review process
Understanding the Basics of Microarrays http://www.ncbi.nlm.nih.gov/About/primer/microarrays.html	This publication from the National Center for Biotechnology Information (NCBI) provides an overview of DNA microarrays explaining gene expression, the technology underlying microarrays, the purpose and importance of microarrays, and the basics of microarray experiments.
Gene Expression Omnibus http://www.ncbi.nlm.nih.gov/geo	GEO: the Gene Expression Omnibus. GEO serves as public repository and online resource for storage and retrieval of gene expression data. GEO currently maintains microarray and serial analysis of gene expression (SAGE) data on over 100
European Bioinformatics Institute http://www.ebi.ac.uk/	The European Bioinformatics Institute (EBI) is a nonprofit organization that focuses on research and services in bioinformatics. EBI's website enables access to gene expression databases (Array Express Archive and Gene Expression Atlas) and microarray analysis tools (Expression Profiler, Next Generation and Bioconductor).
Serial Analysis of Gene Expression Portal http://www.sagenet.org	Sagenet provides a detailed description of serial analysis of gene expression (SAGE). This website also provides SAGE applications, publications, and resources.
Histone Database http://www.research.nhgri.nih.gov/histones	NHGRI histone database Histone sequence information, including posttranslational modifications
Antibody Validation Database http://compbio.med.harvard.edu/antibodies/about	Collect and to share experimental results on antibodies that would otherwise remain in individual laboratories, thus aiding researchers in selection and validation of antibodies.
Chromatin Structure and Function http://www.chromatin.us	Information on chromatin biology, histones and epigenetics (hosted by Jim Bone)

Name and Address	Description
Database for DNA Methylation and Environmental Epigenetic Effects http://www.methdb.de/	Human DNA methylation Database DNA methylation data readily available to public Future develop includes environmental impact on methylation
CpG Island Searcher http://www.uscnorris.com/cpgislands2/cpg.aspx	CpG island searcher CpG Island sequence search algorithm Allows for selection of % methylation and length of (ISLAND?) and gaps between islands
Catalogue of Parent of Origin Effects http://igc.otago.ac.nz/home.html	Imprinted Gene Catalogue Catalogue of parent of origin effects Can search by taxon, chromosome, gene name or key word
Database of Noncoding RNAs http://www.noncode.org	Knowledge database dedicated to ncRNA Information on: class, name, location, related publications, mechanism through which it exerts its function Includes all traditional ncRNAs, but excludes tRNAs and rRNAs
MicroRNA Database http://www.mirbase.org	MicroRNA data resource Searchable database of >16,000 published miRNA sequences and annotation – includes location and sequence of mature miRNA Can search by name, keyword, reference and/or annotation
Epigenome Network of Excellence http://www.epigenome-noe.net	Epigenome Network of Excellence Web site of European interdisciplinary epigenetics research network Includes protocols, an antibody database and reference information on epigenetics
Human Epigenome Project http://www.epigenome.org	The Human Epigenome Project Research Consortium Collaborative effort to catalogue and interpret genome-wide methylation patterns of all human genes and major tissues

Table 2

Online Commercial Resources Used in Manuscript

Name	Address
Applied Biosystems Incorporated	http://www.appliedbiosystems.com
Roche Applied Science	http://www.roche-applied-science.com
Illumina Incorporated	http://www.illumina.com
Affymetrix Incorporated	http://www.affymetrix.com
Millipore	http://www.millipore.com
Sequenom Incorporated	http://www.sequenom.com
Preanalytix	http://www.preanalytix.com
Life Technologies Corporation	http://www.lifetechnologies.com
Ambion	http://www.ambion.com
Qiagen Incorporated	http://www.qiagen.com
DNAGenoteck Incorporated	http://www.dnagenotek.com
Panomics	http://www.panomics.com
Roche Niblegen Incorporated	http://www.nimblegen.com
Arraystar Incorporated	http://www.arraystar.com

APPENDIX E

LICENSE AGREEMENT FOR MANUSCRIPT #1

Springer Publishing Company, Inc. LICENSE TERMS AND CONDITIONS

Mar 31, 2016

This is a License Agreement between University of Pittsburgh -- Theresa Koleck ("You") and Springer Publishing Company, Inc. ("Springer Publishing Company, Inc.") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer Publishing Company, Inc., and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3839380997131
License date	Mar 31, 2016
Licensed content publisher	Springer Publishing Company, Inc.
Licensed content title	Annual Review of Nursing Research, Volume 29 : Genetics
Licensed content date	May 29, 2012
Type of Use	Thesis/Dissertation
Requestor type	Academic institution
Format	Electronic
Portion	chapter/article
Title or numeric reference of the portion(s)	All
Title of the article or chapter the portion is from	Molecular genomic research designs
Editor of portion(s)	Kasper
Author of portion(s)	Baumgartel et al.
Volume of serial or monograph.	29
Issue, if republishing an article from a serial	1
Page range of the portion	None
Publication date of portion	2012
Rights for	Main product
Duration of use	Life of current and all future editions
Creation of copies for the disabled	no
With minor editing privileges	no
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional	no

use	
The lifetime unit quantity of new product	Up to 499
Made available in the following markets	universities
The requesting person/organization is:	Theresa (Timcheck) Koleck
Order reference number	None
Author/Editor	Theresa Koleck
The standard identifier	KoleckETD
Title	Cognitive Function and Breast Cancer: Genomics and Disease Characteristics
Publisher	University of Pittsburgh
Expected publication date	Apr 2017
Estimated size (pages)	250
Total (may include CCC user fee)	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

The following terms are individual to this publisher:

A maximum of 10% of the content may be licensed for republication.
User is responsible for identifying and seeking separate licenses (under this service or otherwise) for, any of such third party materials which are identified anywhere in the work by permission; without a separate license, such third party materials may not be used.

Other Terms and Conditions:

STANDARD TERMS AND CONDITIONS

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.
2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

3. Scope of License; Limitations and Obligations.

3.1 All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order

Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

3.2 General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, Dept 001, P.O. Box 843006, Boston, MA 02284-3006. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3 Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4 In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5 Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6 User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to

the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED “AS IS”. CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. **Miscellaneous.**

8.1 User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

8.2 Use of User-related information collected through the Service is governed by CCC’s

privacy policy, available online here:http://www.copyright.com/content/cc3/en/tools/footer/privacypolicy.html.

8.3 The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

8.4 No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

8.5 The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to info@copyright.com.

v 1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

APPENDIX F

MANUSCRIPT #2: APOLIPOPROTEIN E GENOTYPE AND COGNITIVE FUNCTION IN POSTMENOPAUSAL WOMEN WITH EARLY-STAGE BREAST CANCER

Online Exclusive Article

Apolipoprotein E Genotype and Cognitive Function in Postmenopausal Women With Early-Stage Breast Cancer

Theresa A. Koleck, BSN, RN, Catherine M. Bender, PhD, RN, FAAN, Susan M. Sereika, PhD, Gretchen Ahrendt, MD, Rachel C. Jankowitz, MD, Kandace P. McGuire, MD, Christopher M. Ryan, PhD, and Yvette P. Conley, PhD

Purpose/Objectives: To examine the role of apolipoprotein E (APOE) genotype in the cognitive function of postmenopausal women with early-stage breast cancer prior to initiation of adjuvant therapy and over time with treatment.

Design: Longitudinal, genetic association study.

Setting: Urban university cancer center.

Sample: Three cohorts of postmenopausal women: 37 women with breast cancer receiving chemotherapy and anastrozole, 41 women with breast cancer receiving anastrozole alone, and 50 healthy women.

Methods: Cognitive function was evaluated three times during a 12-month period using a comprehensive neuropsychological test battery. Participants were genotyped and classified based on the presence or absence of at least one APOE ε4 allele. Multiple linear regression was used to determine if APOE genotype accounted for observed variability in cognitive function data.

Main Research Variables: APOE genotype, breast cancer treatment, and cognitive function.

Findings: Performance or changes in performance on tasks of executive function, attention, verbal learning and memory, and visual learning and memory were found to be influenced by APOE genotype and/or interactions between APOE genotype and study cohort.

Conclusions: The results indicate that cognitive function in postmenopausal women with breast cancer is modified by APOE genotype and the combination of APOE genotype and treatment.

Implications for Nursing: APOE genotype, along with other biomarkers, may be used in the future to assist nurses in identifying women with breast cancer most at risk for cognitive decline.

Key Words: breast neoplasms; cognition; genes; biologic markers

ONF, 41(6), E313–E325. doi: 10.1188/14.ONF.E313-E325

Breast cancer is the most prevalent form of cancer, excluding skin cancer, among women in the United States, with an estimated 232,340 new cases of invasive breast cancer and 64,640 new cases of carcinoma in situ diagnosed in 2013 (American Cancer Society, 2013). Fortunately, in the United States, the overall five-year relative survival rate for women with breast cancer, inclusive of all stages, is 89% (Howlader et al., 2011), making women with breast cancer the largest group of cancer survivors in the United States at 2.9 million women (American Cancer Society, 2013). However, survivorship comes with long-term and late effects related to cancer and/or cancer treatment for a large number of breast cancer survivors.

One of the most common and problematic phenomenon experienced by breast cancer survivors is adjuvant therapy-related cognitive decline (Bender et al., 2006; Downie, Mar Fan, Houédé-Tchen, Yi, & Tannock, 2006; Hurria et al., 2006; Jenkins et al., 2006; Mehnert et al., 2007; Schagen et al., 1999; Schilder et al., 2009; Shilling & Jenkins, 2007). A large body of evidence exists to objectively support these reported deficits (Falleti, Sanfilippo, Maruff, Weih, & Phillips, 2005). In addition, growing evidence suggests that women with breast cancer have poorer cognitive function compared to healthy women prior to the initiation of adjuvant therapy (Hermelink et al., 2007; Schilder et al., 2010; Wefel et al., 2004; Wefel, Saleeba, Buzdar, & Meyers, 2010). Even small changes in cognitive function can have a major impact on a survivor's quality of life, affecting relationships with family and friends, educational and career decisions, job performance, emotional state, the ability to make informed treatment decisions, and adherence to cancer therapy (Boykoff, Moieni, & Subramanian, 2009; Munir, Burrows, Yarker, Kalawsky, & Bains, 2010; Myers, 2012; Stilley, Bender, Dunbar-Jacob, Sereika, & Ryan, 2011; Tchen et al., 2003; Von Ah, Habermann, Carpenter, & Schneider, 2013).

However, discrepancies remain in the percentage of women with breast cancer exhibiting cognitive changes, the severity of the change, and the specific cognitive

domains affected (Falleti et al., 2005; Janelsins et al., 2012). It also remains unclear if all women with breast cancer or only a subset of these women are at risk for poorer cognitive function at pretreatment or for cognitive decline with therapy. Therefore, understanding the variability in cognitive changes in women with breast cancer is key to better predict which women are most at risk for poorer pretreatment cognitive function, as well as cognitive decline with adjuvant therapy, and to tailor and personalize interventions to mitigate the effects of cognitive changes for these women.

Potential Mechanisms Related to Cognitive Decline

Oxidative Stress

A potential mechanism to account for the poorer pre-therapy cognitive function and the cognitive changes observed in women with breast cancer is oxidative stress. Oxidative stress has been implicated in other, more severe cognitive conditions including mild cognitive impairment, Parkinson disease, and Alzheimer disease (Bonda et al., 2010; Mariani, Polidori, Cherubini, & Mecocci, 2005). Oxidation refers to the removal of an electron from an atom or molecule and occurs normally in humans as part of mechanisms such as mitochondrial and peroxisomal metabolism, but also can be the result of exogenous exposures to various agents including ultraviolet light, chemotherapeutics, and environmental toxins (Finkel & Holbrook, 2000).

One of the byproducts of oxidation is free radicals. Free radicals that contain oxygen, or reactive oxygen species (ROS), are of particular interest within biologic systems. ROS are positively charged, unstable atoms or molecules that try to achieve stability by taking electrons from other atoms or molecules. This process of stealing electrons can result in cellular and DNA damage along with the creation of additional free radicals, generating a chain reaction of even more damage that can ultimately result in neuronal dysfunction (Finkel & Holbrook, 2000). To combat excessive ROS burden, humans have antioxidant defenses, including specific enzymes, peptides, and vitamins. Therefore, oxidative stress is the sum of ROS production and antioxidant capability for ROS elimination (Azzi, 2007; Finkel & Holbrook, 2000).

The cellular environment of a woman with breast cancer is one of increased oxidative stress. Research has shown that individuals with cancer have higher levels of oxidative stress markers prior to treatment than healthy controls (Amin, Mohamed, El-Wakil, & Ibrahem, 2012; Blasiak et al., 2004; Hamed, Zakhary, & Maximous, 2012). In addition, chemotherapy serves as an exogenous source of ROS (Conroy et al., 2012; Joshi et al., 2005; Kasapovic et al., 2010), and anti-estrogen therapies such as aromatase inhibitors essentially block the production of estrogen, which performs an antioxidant role in the brain (Strehlow et al., 2003; Unfer, Conterato, Da Silva, Duarte, & Emanuelli, 2006). Because of high metabolic demands and low antioxidant capacity, brain cells are particularly vulnerable to oxidative damage. For additional detail on the role of chemotherapy and estrogen in cognitive decline, the authors recommend a review article by Walker, Drew, Antoon, Kalueff, and Beckman (2012).

Considering the role oxidative stress plays in poorer cognitive function, the potential increased oxidative stress influence on the brain cells of women with breast cancer, and the variability seen between women with respect to cognitive changes, exploring genetic underpinnings of this observed variability is logical, starting with candidate genes known to influence and/or modify the response to oxidative stress.

Apolipoprotein E

Evidence suggests that apolipoprotein E (APOE) performs antioxidant activities throughout the body (Hayek, Oiknine, Brook, & Aviram, 1994), in addition to its better known function as a regulatory protein involved in cholesterol and phospholipid metabolism (Mahley, Innerarity, Rall, & Weisgraber, 1984). Three functionally distinct APOE isoforms exist in humans, E2, E3, and E4, which correspond to the three normal variant alleles, ε2, ε3, and ε4, respectively. These allele variants differ from each other at two amino acid sites (Mahley et al., 1984). The antioxidant ability of APOE appears to be isoform-dependent with the E2 isoform having the greatest antioxidant capacity and the E4 isoform having the least antioxidant capacity (i.e., E2 > E3 > E4) (Jolivalt et al., 2000; Miyata & Smith, 1996; Pedersen, Chan, & Mattson, 2000). Additional information about APOE genotype and oxidative stress can be found in Jofre-Monseny, Minihane, and Rimbach (2008).

In addition, a well-established relationship exists between the presence of one or more ε4 alleles and increased risk of Alzheimer disease (Farrer et al., 1997; Richard & Amouyel, 2001; Sadigh-Eteghad, Talebi, & Farhoudi, 2012). Numerous studies also have found a relationship between the ε4 allele and poorer cognitive functioning in healthy middle-aged and older adult populations (Flory, Manuck, Ferrell, Ryan, & Muldoon, 2000; Hofer et al., 2002; Izaks et al., 2011; Wehling, Lundervold, Standnes, Gjerstad, & Reinvang, 2007). However, only one previous study has investigated the association between APOE genotype and cognitive change in women with breast cancer. In this cross-sectional study of 80 long-term breast cancer and lymphoma survivors, who had previously received standard dose chemotherapy and were now an average

of 8.8 years post-treatment, Ahles et al. (2003) found that the presence of at least one ε4 allele was associated with poorer performance in visual memory, spatial ability, and psychomotor functioning compared to survivors who did not possess an ε4 allele. However, the interpretations of these findings are limited by the lack of pretreatment data, longitudinal assessment, and healthy control group for comparison. In addition, the substantial length of time post-treatment does not inform the immediate effects of APOE genotype and treatment on cognitive function.

Therefore, because of the presumed increase in oxidative stress from cancer, chemotherapy, and antiestrogen therapy, combined with the known impact of oxidative stress on cognitive function and the variability in antioxidant capacity by APOE isoform, the purpose of the current study was to explore the role of APOE genotype in the cognitive function of postmenopausal women with early-stage breast cancer prior to the initiation of adjuvant chemotherapy and/or antiestrogen therapy and over time through the first year of adjuvant treatment.

Methods

Participants and Setting

Participants were recruited for this exploratory, genetic ancillary study from the Anastrozole Use in Menopausal Women (AIM) study (R01 CA107408), a longitudinal prospective cohort study investigating the impact of the anti-estrogen therapy, anastrozole, on changes in cognitive function in postmenopausal women with breast cancer. The final sample for this ancillary study (N = 128) was comprised of three cohorts of postmenopausal women: (a) women with breast cancer who received chemotherapy plus anastrozole (n = 37), (b) women with breast cancer who received anastrozole alone (n = 41), and (c) healthy, control women matched on age and years of education to the participants with breast cancer (n = 50).

Women with breast cancer were recruited from the Comprehensive Breast Cancer Program of the University of Pittsburgh Cancer Institute. Healthy women were recruited using a variety of approaches including referral from women in the breast cancer cohorts, advertisements, and random digit dialing through the University Center for Social and Urban Research.

Participants currently undergoing data collection for the AIM study were simultaneously recruited to obtain a genetic sample for the ancillary study. Participants who previously completed data collection for the AIM study, and gave permission to be recontacted, were contacted for the purpose of procuring a genetic sample. Both the AIM study and ancillary study were approved by the University of Pittsburgh Institutional Review Board. Informed consent was obtained from all study participants for the parent study and the ancillary genetic study.

Inclusion criteria for all participants include being postmenopausal, having a maximum age of 75 years, having the ability to speak and read English, and completion of a minimum of eight years of education. An additional inclusion criterion for women with breast cancer was newly diagnosed early-stage breast cancer (i.e., stages I, II, or IIIa) based on the Tumor, Nodes, Metastasis (TNM) Classification of Malignant Tumors (Edge et al., 2010). Exclusion criteria for all participants include self-reported hospitalization for a psychiatric illness within the past two years and a history of neurologic illness or cancer. In addition, women with breast cancer with clinical evidence of distant metastases were deemed ineligible.

Evaluation of Cognitive Function

Cognitive function was evaluated at three time points in all study participants. For women with breast cancer receiving chemotherapy plus anastrozole, cognitive function was assessed after primary surgery but prior to the initiation of chemotherapy (T0), prior to the initiation of anastrozole (T1), and six months after the initiation of anastrozole (T2). Cognitive function was

Table 1. Neuropsychological Tests According to Cognitive Function Factors

Factor	Neuropsychological Test
Attention	CANTAB Spatial Working Memory Test (Owen et al., 1995) CANTAB Stockings of Cambridge Test (Owen et al., 1995) Digit Vigilance Test (Matthews, 1964)
Executive function	Delis Kaplan Color Word Interference Test (Delis et al., 2001) Verbal Fluency Test (Delis et al., 2001) Trail Making Test B (Reitan & Wolfson, 1985)
Psychomotor efficiency	Grooved Pegboard Test (Matthews, 1964) Digit Symbol Substitution Test (Wechsler, 1981)
Verbal learning and memory	Rivermead Behavioral Memory Test (Wilson et al., 1989) Rey Auditory Verbal Learning Test (Rey, 1964)
Visual learning and memory	CANTAB Paired Associates Learning Test (Owen et al., 1995) Rey Complex Figure Test (Osterrieth, 1944)
Visuospatial ability	CANTAB Rapid Visual Information Processing Test (Owen et al., 1995)

CANTAB—Cambridge Neuropsychological Test Automated Battery

Knowledge Translation

Possession of one or more apolipoprotein E (APOE) ε4 alleles has been associated with decreased antioxidant capacity and increased risk of Alzheimer disease.

Variability in APOE genotype may partially explain observed variation in cognitive changes in women with and receiving treatment for breast cancer.

Potential modifications of cancer- and treatment-related cognitive changes in women with breast cancer by genetic variation should be further investigated.

evaluated in women who received anastrozole alone prior to the initiation of anastrozole (T0), six months after the initiation of anastrozole (T1), and 12 months after the initiation of anastrozole (T2). Healthy controls were assessed at comparable time points: baseline (T0), six months after T0 (T1), and 12 months after T0 (T2).

Cognitive function was measured using a comprehensive battery of neuropsychological tests encompassing six cognitive domains: attention, learning and memory, psychomotor speed, mental flexibility, executive function, and visuospatial ability. Neuropsychological tests were selected based on test validity, reliability, and sensitivity, as well as on the availability of alternative, equivalent forms to minimize the influence of practice effects. The battery was administered to study participants by research nurses trained by a clinical neuropsychologist. The average time for completion was 90 minutes. The neuropsychological tests comprising the battery and the reduction of the dimensionality of the cognitive function data has been described in detail previously (Bender et al., 2013). The six resulting composite cognitive function factors and the neuropsychological tests comprising each factor are detailed in Table 1. All cognitive measures have been demonstrated to be sensitive to changes in cognitive function in women with breast cancer (Bender et al., 2010).

Covariates and Confounders

Age and estimated verbal intelligence (National Adult Reading Test-Revised) (Nelson, 1981) were measured at T0. Time-dependent covariates including depression (Beck Depression Inventory–II) (Beck, Steer, & Brown, 1996), anxiety (Profile of Mood States [POMS] tension-anxiety subscale) (McNair, Lorr, & Droppleman, 1992), fatigue (POMS fatigue-inertia subscale) (McNair et al., 1992), and pain (Brief Pain Inventory) (Cleeland, 1989) were assessed at each time point.

Genotyping for Apolipoprotein E

A sample of 3 cc of whole blood or 2 cc of saliva was collected from each participant. DNA was extracted from peripheral leukocytes using a simple salting out procedure (Miller, Dykes, & Polesky, 1988) or from saliva using the protocol and reagents supplied with the Oragene DNA collection kits (DNA Genotek, 2012). Genotypes were determined for the two functional single nucleotide polymorphisms (SNPs) for the APOE gene, *rs429358* and *rs7412*, that comprise the ε2, ε3, and ε4 alleles. Genotype for *rs429358* was determined via TaqMan® allelic discrimination, and genotype for *rs7412* was determined by inclusion in an i-PLEX® MassARRAY® multiplex assay. Positive and negative controls were included. Genotype data were double blind culled by two individuals, and discrepancies were rectified by review of raw data. SNP genotypes for *rs429358* and *rs7412* were combined for each participant, as detailed in Table 2, to determine APOE genotype. Participant genotypes were then classified based on the presence (i.e., ε4/ε4, ε2/ε4, and ε3/ε4) or absence (i.e., ε2/ε2, ε2/ε3, and ε3/ε3) of one or more APOE ε4 alleles.

Statistical Analysis

The statistical analysis was carried out using StataSE®, version 12. A detailed descriptive analysis of all data, including demographic data, was initially performed. Data were screened for all assumptions required for the planned linear regression analysis (e.g., linearity, normality), and sources of missing data were investigated. The comparability of baseline covariate and confounder data and baseline cognitive ability between participants included in the ancillary analysis and remaining participants from the parent study was assessed using independent t tests to evaluate equality of means. In addition, the comparability of demographic and baseline covariate and confounder data among APOE ε4 status and study cohorts was assessed using analysis of variance and Pearson's chi-square tests of independence.

Multiple linear regression was used to investigate the effect of APOE genotype on all six cognitive factors, both cross-sectionally for each time point (i.e., T0,

Table 2. APOE Genotype Determination

APOE Genotype	*rs429358* Allele	*rs7412* Allele
ε2/ε2	T	T
ε3/ε3	T	C
ε2/ε3	T	CT
ε2/ε4	CT	CT
ε3/ε4	CT	C
ε4/ε4	C	C

APOE—apolipoprotein E

T1, and T2) and longitudinally using change scores (i.e., T0–T1, T0–T2, and T1–T2). To obtain minimally confounded estimates of effect, all evaluated predictors were included in each model. Age, estimated intelligence, and study cohort were incorporated as fixed covariates and confounders. Time-dependent covariates and confounders (i.e., depression, anxiety, fatigue, and pain scores) for a particular assessment time point, or the change in a time-dependent covariate and confounder from assessment to assessment, were incorporated into each model as appropriate. Because the authors were interested in how the effect of APOE genotype on cognitive function may be modified by the prescribed treatment regimen, interactions between APOE ε4 absence or presence and study cohort were initially examined. If no significant interactions were observed, a main effects model, considering only APOE ε4 absence/presence and study cohort, was fit for each cognitive function factor. Women with no ε4 alleles and the healthy control cohort served as the reference groups in the regression analysis. Unstandardized regression coefficients and significance tests at a two-tailed significance level of 0.05 were used to determine if APOE ε4 genotype status or APOE ε4 genotype by study cohort interactions improved model fit and, therefore, account for observed variability in the cognitive function data.

Table 3. Comparison of T0 Characteristics of AIM Study Participants Included and Not Included in the APOE Analysis (N = 366)

	Included (n = 128)		Not Included (n = 238)		
Characteristic	$\bar{X}$	SD	$\bar{X}$	SD	p[a]
Age (years)	59.31	5.699	60.66	6.432	0.048*
Years of education	15.22	3.157	14.55	2.66	0.032*
Estimated intelligence[b]	110.25	9.184	108.33	9.149	0.057
Depression[c]	4.8	5.161	6.1	6.608	0.055
Anxiety[d]	7.64	5.698	7.59	5.784	0.931
Fatigue[e]	5.2	5.77	5.67	5.575	0.456
Pain[f]	1.25	1.98	1.51	2.292	0.262
Visual learning and memory[g]	0.107	0.785	–0.1139	0.839	0.015*
Executive function[g]	0.1316	0.598	0.0827	0.707	0.506
Verbal learning and memory[g]	–0.0591	0.843	–0.2237	0.809	0.068
Attention[g]	–0.1119	0.695	–0.2652	0.739	0.054
Psychomotor efficiency[g]	0.0558	0.738	–0.1555	0.829	0.016*
Visuospatial ability[g]	–0.0902	1.018	–0.0847	0.899	0.958

* $p < 0.05$
[a] Independent t tests were used to compare means between AIM study participants included and not included in the APOE analysis.
[b] National Adult Reading Test–Revised verbal IQ score
[c] Beck Depression Inventory–II
[d] Profile of Mood States tension/anxiety subscale
[e] Profile of Mood States fatigue/inertia subscale
[f] Brief Pain Inventory Pain Right Now score
[g] Z score
AIM—Anastrozole Use in Menopausal Women; APOE—apolipoprotein E

For each regression model, the authors examined the residuals to identify any sources of model misspecification or outliers and influential observations that may have impacted the validity of the regression findings. The screening of residuals identified several models that did not meet normality or homogeneous variance assumptions and/or contained ill-fitted observations. In cases of nonnormality or heterogeneous variance, a series of data transformations were conducted in an attempt to induce normality and homoscedasticity. To evaluate the *robustness* of findings, a regression model excluding points determined to be influential, as well as a robust regression model using Huber and biweight iterations, was generated. Models eliminating potentially influential multivariate-outlier cases or diminishing the weight of potentially influential univariate-outlier cases were created, as needed, to conclude the sensitivity analysis. Unstandardized regression coefficients, p values, and the correlations of fitted values were compared between the models.

Findings

Genetic samples were collected from 137 (37%) of the 366 participants from the AIM parent study. Of the 137, 5 participants (4%) had indeterminable genotypes and 4 participants (3%) had incomplete cognitive function or covariate and confounder information at T0. The women included in the APOE analysis (n = 128) were marginally younger (p = 0.048) and better educated (p = 0.032) than AIM study participants not included in the APOE analysis (n = 238) (see Table 3). Women in the APOE analysis also had higher unadjusted mean baseline visual learning and memory (p = 0.015) and

Table 4. Sample Characteristics (N = 128)

Characteristic	Chemotherapy Plus Anastrozole (n = 37)				Anastrozole Alone (n = 41)				Healthy Controls (n = 50)				
	ε4 (n = 11)		No ε4 (n = 26)		ε4 (n = 9)		No ε4 (n = 32)		ε4 (n = 16)		No ε4 (n = 34)		
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	p[a]
Age (years)	58.64	4.61	58.5	5.666	61.56	4.613	61.03	5.608	60.25	6.557	57.5	5.594	0.112
Years of education	16.27	3.349	15.42	2.595	15.67	2.958	14.97	3.605	15	2.633	14.94	3.428	0.847
Estimated intelligence (NART)	110.58	7.299	108.2	9.754	109.22	6.378	105.92	9.516	115.368	9.373	113.52	7.375	0.002*
T0 depression (BDI-II)	3.09	2.625	5.85	5.655	3	2.291	5.34	6.378	4.06	4.669	4.88	4.848	0.548
T0 anxiety (POMS tension/anxiety)	10.64	7.514	9.45	5.896	6.44	3.245	7.16	4.712	6.13	5.084	6.79	6.188	0.145
T0 fatigue (POMS fatigue/inertia)	3.91	4.061	6.23	6.855	2.78	3.346	5.69	6.188	5	6.501	5.12	5.139	0.674
T0 pain (BPI Right Now)	0.91	1.375	1.69	2.112	1.56	1.74	1.28	1.955	0.63	1.544	1.21	2.307	0.638

Characteristic	n	n	n	n	n	n	p[a]
Married	7	20	5	24	8	22	0.433
Children	8	23	7	27	13	31	0.678
Caucasian	10	25	9	32	15	33	0.672
Cancer stage							
• I	8	11	8	25	–	–	–
• IIa	3	10	–	7	–	–	
• IIb	–	3	1	–	–	–	
• IIIa	–	2	–	–	–	–	

* $p < 0.05$

[a] One-way analysis of variance was used to compare study cohort means of continuous variables. Pearson's chi-square tests of independence was used to examine the general associations between categorical variables.

BDI-II—Beck Depression Inventory–II; BPI—Brief Pain Inventory; NART—National Adult Reading Test–Revised verbal IQ score; POMS—Profile of Mood States

Note. At baseline, participants were not experiencing depression, anxiety, fatigue, or pain. Although not significant, women in the chemotherapy plus anastrozole group had somewhat higher anxiety scores at baseline.

psychomotor efficiency ($p = 0.016$) factor z scores than the remaining AIM study participants. No relationship was observed between study cohort and ε4 genotype status ($\chi^2 = 1.192$, $p = 0.551$). Study cohort by ε4 absence/presence groups differed slightly on estimated intelligence ($p = 0.002$) (see Table 4). The study cohorts did not differ on age, years of education, or baseline levels of depression, anxiety, fatigue, and pain. In general, study participants were Caucasian (97%), married (67%), and had one or more child (85%).

Cross-Sectional Time Point Analysis

Significant time point analysis findings are summarized in Table 5. The time point analysis indicated that possession of one or more ε4 alleles contributes to poorer verbal learning and memory performance at T0 ($\beta = -0.334$, $p = 0.031$) and T1 ($\beta = -0.3222$, $p = 0.038$) regardless of cancer or treatment status. Although not statistically significant, this trend extends to T2 ($\beta = -0.2891$, $p = 0.064$). The combination of anastrozole-alone group membership and possession of one or more ε4 alleles contributes negatively to executive function performance both at T0 ($\beta = -0.4448$, $p = 0.088$) and T1 ($\beta = -0.5771$, $p = 0.033$) (see Figure 1).

Longitudinal Change Score Analysis

Significant change score analysis findings are summarized in Table 6. The change score analysis revealed a significant decline in visual learning and memory from T1 to T2 ($\beta = -0.269$, $p = 0.027$) for women with one or more ε4

alleles compared to women with no ε4 alleles regardless of cancer or treatment status. In addition, the combination of anastrozole-alone group membership and possession of one or more ε4 alleles negatively impacted change in visual learning and memory from T0 to T2 (β = –0.567, p = 0.042) (see Figure 2). The combination of anastrozole-alone group member and possession of one or more ε4 alleles contributes negatively to the change in attention from T1 to T2 (β = –0.5715, p = 0.045) (see Figure 3). In addition, the combination of chemotherapy plus anastrozole group membership and possession of one or more ε4 alleles had a positive impact on verbal learning and memory scores from T0 to T2 (β = 0.5468, p = 0.064) (see Figure 4).

Table 5. Cognitive Factors With Significant Cross-Sectional Assessment Results

	APOE ε4 Presence		APOE ε4 Presence by Chemotherapy Plus Anastrozole Interaction		APOE ε4 Presence by Anastrozole Alone Interaction	
Time and Model	**β**	**p**	**β**	**p**	**β**	**p**
Executive Function						
T0 (n = 128)						
Interaction	0.2675	0.102	–0.0654	0.795	–0.4448	0.088*
Main effects	0.1257	0.244				
T1 (n = 125)						
Interaction	0.3292	0.047*	–0.2429	0.353	–0.5771	0.033*
Main effects	0.1047	0.341				
T2 (n = 112)						
Interaction	0.1237	0.537	0.086	0.793	–0.3331	0.323
Main effects	0.0679	0.62				
Verbal Learning and Memory						
T0 (n = 128)						
Interaction	–0.0522	0.882	–0.633	0.079	–0.3464	0.349
Main effects	–0.334	0.031*				
T1 (n = 125)						
Interaction	–0.0899	0.417	–0.0993	0.789	–0.3895	0.309
Main effects	–0.3222	0.038*				
T2 (n = 112)						
Interaction	–0.1778	0.436	–0.3244	0.384	–0.0774	0.84
Main effects	–0.2891	0.064*				

* $p < 0.1$; estimates controlled for age, estimated intelligence, depression, anxiety, fatigue, and pain scores.

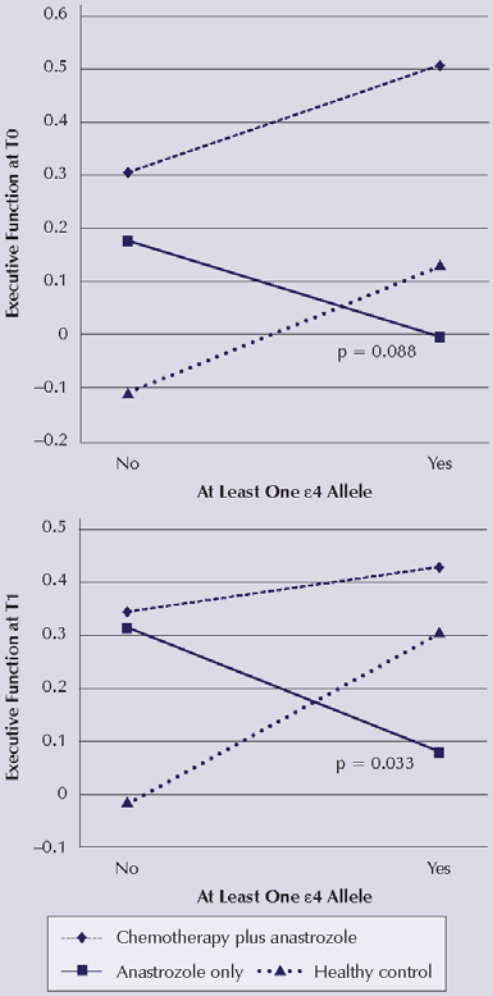
APOE—apolipoprotein E

Note. The healthy control cohort and women with no ε4 alleles served as the reference groups in the analysis.

Discussion

This exploratory study investigated the role of APOE genotype in cognitive function of postmenopausal women with early-stage breast cancer and represents the first study to examine the effect of APOE genotype, breast cancer, and breast cancer treatment simultaneously on cognitive function over time. In the individual time point analysis, the authors found significant or moderately significant associations between the possession of one or more ε4 alleles and poorer verbal learning and memory performance, regardless of cancer or treatment status, at all three assessment time points. Study cohort by ε4 status interactions also were observed at baseline and at the first post-treatment assessment time point for the executive function factor, with the combination of anastrozole-alone group membership and possession of one or more ε4 alleles contributing to poorer performance on executive function tasks. When the authors assessed the effect of possession of one or more ε4 alleles on changes in cognitive function over time, a significant main effect was found that was indicative of a decrease in visual learning and memory performance from T1–T2, regardless of cancer or treatment status, as well as two significant interaction effects. Specifically, anastrozole-alone group membership in combination with ε4 carrier status contributed to a decrease in attention scores from the first post-treatment (six months post-anastrozole initiation) to the second post-treatment assessment (12 months post-anastrozole initiation), and chemotherapy plus anastrozole group membership in combination with ε4 carrier status contributed to an improvement in verbal learning and memory from baseline to the second post-treatment assessment.

Consistent with findings previously reported in the literature on the relationship between APOE genotype and memory in the general adult population, the authors found that possession of one or more ε4 alleles was associated with poorer verbal learning and memory performance across all study participants, regardless of study cohort or treatment status, at every assessment time point (Caselli et al., 2011; Flory et al., 2000; Hofer et al., 2002; Nilsson, Nyberg, & Bäckman, 2002; Wehling et al., 2007). The authors propose that the marginally significant findings observed at T2 could be a reflection of practice effects (Lezak, Howieson, & Loring, 2004).



Note. Mean Z scores were calculated for each apolipoprotein E ε4 status and treatment combination based on mean covariate and confounder values. P values for the significant or marginally significant interactions are displayed in each graph.

Figure 1. Mean Z Scores for Interaction Effects: Executive Function

Executive function was the other cognitive factor found to have significant cross-sectional APOE genotype effects. Of note, while the main effect β coefficient contributes positively to the model for all participants, the interaction β coefficient contributes negatively to the model, nullifying the main effect and contributing an overall negative input to the baseline executive function performance for women prescribed anastrozole possessing one or more ε4 alleles. This latter finding, in particular, not only adds to the literature supporting the notion that women with breast cancer have poorer cognitive function prior to the initiation of adjuvant therapy compared to healthy controls, but also extends the knowledge, suggesting that cognitive changes are potentially augmented by genetic variation and the biologic characteristics of a woman's breast cancer that determine treatment regimens (Ahles & Saykin, 2007; Vardy, Wefel, Ahles, Tannock, & Schagen, 2008). A similar finding was observed at the first post-treatment assessment, lending support to the proposed increased oxidative stress hypothesis; however, this trend did not significantly extend to the second post-treatment assessment.

Of note, the authors found that a chemotherapy plus anastrozole treatment regimen in combination with possession of one or more ε4 alleles actually positively contributed to verbal learning and memory performance from baseline to the second post-treatment assessment; this same trend is observed for anastrozole treatment regimen in combination with ε4 carrier status. Although unexpected based on the proposed oxidative stress hypothesis, which postulates that women with breast cancer receiving chemotherapy (i.e., highest amount of oxidative stress) who also possessed one or more ε4 alleles (i.e., least antioxidant capacity) would experience the greatest cognitive decline, this result is not entirely unfounded. In fact, evidence suggests that possession of one or more ε4 alleles may be cognitively advantageous early in life (Hubacek et al., 2001; Yu, Lin, Chen, Hong, & Tsai, 2000). Mondadori et al. (2007) found the ε4 allele to be associated with better episodic memory performance when compared to ε2 and ε3 alleles in healthy, young ($\bar{X}$ age = 22.8 years, SD = 4) adults. In addition, results from the functional magnetic resonance imaging component of the study suggest that the ε4 allele is associated with more economic use of neural learning resources (Mondadori et al., 2007). Several studies considering the effect of the ε4 allele in healthy, middle-aged adults report minimal if any difference in cognitive function performance between heterozygous ε4 carriers and noncarriers (Han & Bondi, 2008; Izaks et al., 2011; Jorm et al., 2007); however, although comparable in neuropsychological task performance, cognitively intact middle- and older-aged ε4 carriers demonstrate greater brain activity during learning and memory tests than their matched ε3 counterparts (Bondi, Houston, Eyler, & Brown, 2005; Wishart et al., 2006). Therefore, this unanticipated longitudinal improvement may be partially accounted for by an undefined protective function of the ε4 allele, more efficient learning (i.e., practice effects), and an increased magnitude and extent of neural resource use by the chemotherapy plus anastrozole cohort on verbal learning and memory tasks. As the current study did not incorporate brain imaging, the two latter hypotheses

Table 6. Cognitive Factors With Significant Longitudinal Change Score Results

	APOE ε4 Presence		APOE ε4 Presence by Chemotherapy Plus Anastrozole Interaction		APOE ε4 Presence by Anastrozole Alone Interaction	
Time and Model	**β**	**p**	**β**	**p**	**β**	**p**
Visual Learning and Memory						
T0–T1 (n = 124)						
Interaction	0.1375	0.371	0.209	0.402	–0.1525	0.548
Main effects	0.154	0.133				
T0–T2 (n = 112)						
Interaction	0.0498	0.76	0.1082	0.681	–0.567	0.042*
Main effects	–0.0604	0.592				
T1–T2 (n = 111)						
Interaction	–0.087	0.622	–0.1782	0.542	–0.5112	0.088*
Main effects	–0.269	0.027*				
Verbal Learning and Memory						
T0 –T1 (n = 124)						
Interaction	–0.0651	0.722	0.4485	0.133	–0.0911	0.763
Main effects	0.0347	0.777				
T0–T2 (n = 112)						
Interaction	–0.1261	0.486	0.5468	0.064*	0.1539	0.616
Main effects	0.0717	0.562				
T1–T2 (n = 111)						
Interaction	–0.0428	0.811	0.053	0.857	0.1105	0.713
Main effects	0.0005	0.997				
Attention						
T0 –T1 (n = 124)						
Interaction	0.0409	0.785	–0.258	0.29	0.1385	0.576
Main effects	0.0069	0.945				
T0–T2 (n = 112)						
Interaction	0.1523	0.375	–0.2949	0.289	–0.3997	0.171
Main effects	–0.0336	0.773				
T1–T2 (n = 111)						
Interaction	0.1539	0.359	–0.1669	0.547	–0.5715	0.045*
Main effects	–0.0408	0.722				

* $p < 0.1$; estimates controlled for age, estimated intelligence, depression, anxiety, fatigue, and pain change scores.

APOE—apolipoprotein E

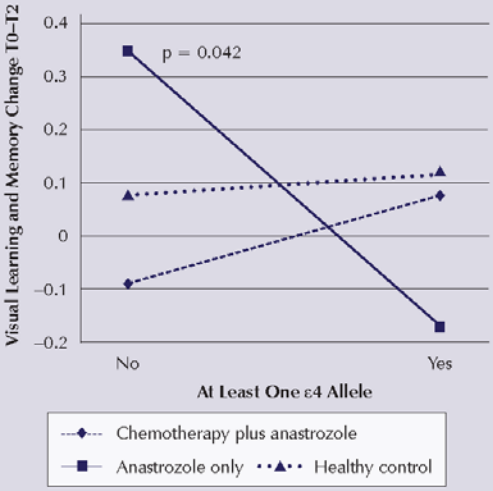
Note. The healthy control cohort and women with no ε4 alleles served as the reference groups in the analysis.

could not be explored. Alternatively, treatment of the underlying cancer (of which cancers prescribed chemotherapy and anastrozole are more aggressive) may result in improvement of symptoms, including cognitive function, over time.

To the authors' knowledge, only one study has previously examined the effect of APOE genotype on cognitive function in individuals with breast cancer. Ahles et al. (2003) reported significantly poorer performance on tasks of visual memory, spatial ability, and psychomotor functioning in long-term breast cancer and lymphoma survivors treated with chemotherapy with one or more ε4 alleles compared to those with no ε4 alleles. The results from Ahles et al. (2003) are difficult to compare to the current study because of the use of a cross-sectional design, the focus on long-term ($\bar{X}$ = 8.8 years post-treatment) cognitive functioning, inclusion of lymphoma survivors, and inability to examine treatment effects. One other study has explored genetic modification of cancer- and therapy-related cognitive changes in women with breast cancer. Small et al. (2011) investigated the influence of catechol-O-methyltransferase (COMT) genotype on cognitive performance six months after completion of treatment in women with breast cancer who received (a) chemotherapy with or without radiotherapy or (b) radiotherapy only and (c) healthy controls with no history of cancer. The results of the study indicated that COMT valine carriers treated with chemotherapy performed more poorly on tasks of attention than healthy controls who were also valine carriers. The results from these studies and the current study all provide evidence for the modification of cancer- and treatment-related cognitive changes in women with breast cancer by genetic variation.

Limitations

Although the results of this exploratory study are informative, a number of limitations should be acknowledged. First, the study sample size was relatively small, limiting the authors' ability to detect small and moderate effects; however, the findings from this study can be used to obtain more accurate sample size estimations for future investigations. The small sample size also did not allow the authors to evaluate dose-response relationships among heterozygous ε4 carriers and homozygous (ε4, ε4) individuals. Second, the



Note. Mean Z scores were calculated for each apolipoprotein E ε4 status and treatment combination based on mean covariate and confounder values. P values for the significant or marginally significant interactions are displayed in each graph.

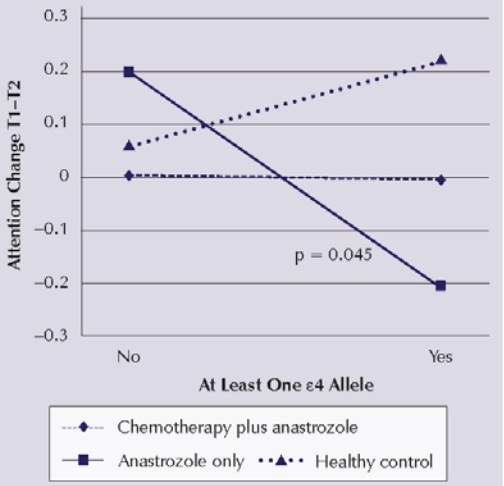
Figure 2. Mean Z Scores for Interaction Effects: Visual Learning and Memory Change

sample was primarily comprised of Caucasian women. The extent to which the results generalize to more diverse populations is unknown. Third, the results indicate that women included in the APOE analysis may be different than those in the AIM study who were not part of the APOE analysis subset. Of little concern are the differences in age and years of education. Although statistically significant, the mean differences in age ($\bar{X}$ = 59.31, SD = 5.699 years for women in the APOE subset versus $\bar{X}$ = 60.66, SD = 6.432 years for those not in the subset) and years of education ($\bar{X}$ = 15.22, SD = 3.157 years for women in the APOE subset versus $\bar{X}$ = 14.55, SD = 2.66 years for those not in the subset) are most likely not clinically meaningful. In contrast, the differences in mean baseline visual learning and memory and psychomotor efficiency z scores, with women in the APOE analysis subset displaying significantly better performance in both factors, may have implications for the validity and generalizability of results. An additional limitation of this study, inherent to all studies that recruit patients with breast cancer following primary surgery, is the potential effects of surgery and stress of cancer diagnosis on cognitive function. Finally, APOE genotype represents only a single insight by which cognitive changes could be augmented in women with breast cancer; additional genes and mechanisms should be considered in the future. However, the authors also would like to acknowledge this study's many strengths, including hypothesis-driven gene selection, pre-adjuvant therapy assessment, longitudinal follow-up, inclusion of a healthy control reference group, evaluation of treatment effects (i.e., chemotherapy and anti-estrogen therapy), and control for many known covariates and confounders of cognitive function.

Conclusions and Implications for Practice and Research

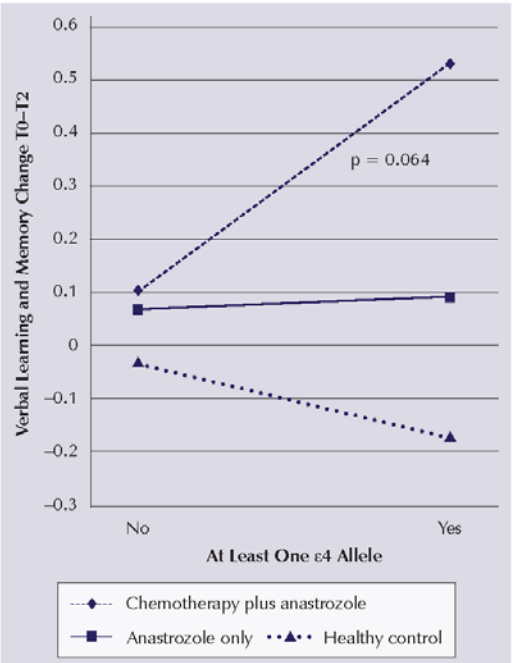
Information gained from the current study adds to the base of knowledge regarding the influence of genetic determinants on poorer cognitive performance and cognitive decline experienced by many survivors of early-stage breast cancer. Although not clinically useful at this point in time, the results from this exploratory analysis indicate modification of cognitive function performance and of cognitive changes over time by both APOE genotype and the combination of APOE genotype and prescribed treatment. In particular, performance on tasks of executive function, attention, verbal learning and memory, and visual learning and memory were influenced by APOE genotype.

Additional research is needed on this topic to further elucidate the role of APOE genotype in cognitive function of women with breast cancer, both in terms of vulnerability to and protection from cognitive decline. The results from this study need to be confirmed in a larger,



Note. Mean Z scores were calculated for each apolipoprotein E ε4 status and treatment combination based on mean covariate and confounder values. P values for the significant or marginally significant interactions are displayed in each graph.

Figure 3. Mean Z Scores for Interaction Effects: Attention Change



Note. Mean Z scores were calculated for each apolipoprotein E ε4 status and treatment combination based on mean covariate and confounder values. P values for the significant or marginally significant interactions are displayed in each graph.

Figure 4. Mean Z Scores for Interaction Effects: Verbal Learning and Memory Change

more diverse sample with similarly detailed pretreatment and longitudinal cognitive function and covariate/confounder assessment. Mechanistic structural and functional brain imaging studies should be conducted to evaluate changes and differences in brain morphology and activation patterns by genotype (Vardy et al., 2008). The functions of oxidative stress and antioxidant capacity on cognitive function in women with breast cancer warrant further investigation as well. Information garnered from future studies will permit a greater understanding of the influence of APOE genotype on cognitive function in women with and receiving treatment for breast cancer, provide the basis for development of biomarkers to identify women most at risk for cognitive changes, and inform novel treatments for women experiencing cognitive decline.

Theresa A. Koleck, BSN, RN, is a doctoral student and Catherine M. Bender, PhD, RN, FAAN, is a professor and doctoral program director in the School of Nursing; Susan M. Sereika, PhD, is a professor in the School of Nursing and a director of the Center for Research and Evaluation and a professor in the Departments of Biostatistics and in the Department of Epidemiology in the Graduate School of Public Health; Gretchen Ahrendt, MD, is an associate professor in the Department of Surgery, Division of Surgical Oncology, in the School of Medicine; Rachel C. Jankowitz, MD, is an assistant professor in the Department of Medicine in the School of Medicine; Kandace P. McGuire, MD, is an assistant professor in the Department of Surgery in the School of Medicine; Christopher M. Ryan, PhD, is a professor in the School of Nursing and in the Department of Psychiatry in the School of Medicine; and Yvette P. Conley, PhD, is a professor in the School of Nursing and in the Department of Human Genetics in the Graduate School of Public Health, all at the University of Pittsburgh in Pennsylvania. This research was supported, in part, by the Genomics of Cognitive Function in Breast Cancer grant from the ONS Foundation, the Targeted Research and Academic Training for Nurses in Genomics training program grant (No. T32NR009759), and the Anastrozole Use in Menopausal Women study grant (No. R01CA107408). Koleck can be reached at tat30@pitt.edu, with copy to editor at ONFEditor@ons.org. (Submitted January 2014. Accepted for publication May 21, 2014.)

References

Ahles, T.A., & Saykin, A.J. (2007). Candidate mechanisms for chemotherapy-induced cognitive changes. *Nature Reviews. Cancer, 7*, 192–201. doi:10.1038/nrc2073

Ahles, T.A., Saykin, A.J., Noll, W.W., Furstenberg, C.T., Guerin, S., Cole, B., & Mott, L.A. (2003). The relationship of APOE genotype to neuropsychological performance in long-term cancer survivors treated with standard dose chemotherapy. *Psycho-Oncology, 12*, 612–619. doi:10.1002/pon.742

American Cancer Society. (2013). Breast cancer: What are the key statistics about breast cancer? Retrieved from http://bit.ly/1sicnbs

Amin, K.A., Mohamed, B.M., El-Wakil, M.A.M., & Ibrahem, S.O. (2012). Impact of breast cancer and combination chemotherapy on oxidative stress, hepatic and cardiac markers. *Journal of Breast Cancer, 15*, 306–312. doi:10.4048/jbc.2012.15.3.306

Azzi, A. (2007). Oxidative stress: A dead end or a laboratory hypothesis? *Biochemical and Biophysical Research Communications, 362*, 230–232. doi:10.1016/j.bbrc.2007.07.124

Beck, A.T., Steer, R.A., & Brown, G.K. (1996). *Beck Depression Inventory–II*. San Antonio, TX: The Psychological Corporation.

Bender, C.M., Sereika, S.M., Berga, S.L., Vogel, V.G., Brufsky, A.M., Paraska, K.K., & Ryan, C.M. (2006). Cognitive impairment associated with adjuvant therapy in breast cancer. *Psycho-Oncology, 15*, 422–430. doi:10.1002/pon.964

Bender, C.M., Sereika, S.M., Houze, M.P., Brufsky, A.M., Berga, S.L., Richey, S.M., . . . Ryan, C.M. (2010). *Deterioration in cognitive function with anastrozole therapy in women with breast cancer.* Presented at the International Cognition in Cancer Task Force Meeting, New York, NY.

Bender, C.M., Sereika, S.M., Ryan, C.M., Brufsky, A.M., Puhalla, S., & Berga, S.L. (2013). Does lifetime exposure to hormones predict pretreatment cognitive function in women before adjuvant therapy for breast cancer? *Menopause, 20*(9), 1–8. doi:10.1097/gme.0b013e3182843eff

Blasiak, J., Arabski, M., Krupa, R., Wozniak, K., Rykala, J., Kolacinska, A., . . . Zadrozny, M. (2004). Basal, oxidative and alkylative DNA damage, DNA repair efficacy and mutagen sensitivity in breast cancer. *Mutation Research, 554*(1–2), 139–148. doi:10.1016/j.mrfmmm.2004.04.001

Bonda, D.J., Wang, X., Perry, G., Nunomura, A., Tabaton, M., Zhu, X., & Smith, M.A. (2010). Oxidative stress in Alzheimer disease:

A possibility for prevention. *Neuropharmacology, 59*(4–5), 290–294. doi:10.1016/j.neuropharm.2010.04.005

Bondi, M.W., Houston, W.S., Eyler, L.T., & Brown, G.G. (2005). fMRI evidence of compensatory mechanisms in older adults at genetic risk for Alzheimer disease. *Neurology, 64*, 501–508. doi:10.1212/01.WNL.0000150885.00929.7E

Boykoff, N., Moieni, M., & Subramanian, S.K. (2009). Confronting chemobrain: An in-depth look at survivors' reports of impact on work, social networks, and health care response. *Journal of Cancer Survivorship: Research and Practice, 3*, 223–232. doi:10.1007/s11764-009-0098-x

Caselli, R.J., Dueck, A.C., Locke, D.E.C., Sabbagh, M.N., Ahern, G.L., Rapcsak, S.Z., . . . Reiman, E.M. (2011). Cerebrovascular risk factors and preclinical memory decline in healthy APOE ε4 homozygotes. *Neurology, 76*, 1078–1084. doi:10.1212/WNL.0b013e318211c3ae

Cleeland, C.S. (1989). Measurement of pain by subjective report. In C.R. Chapman & J.D. Loeser (Eds.), *Advances in pain research and therapy* (vol. 12., pp. 391–403). New York, NY: Raven Press.

Conroy, S.K., McDonald, B.C., Smith, D.J., Moser, L.R., West, J.D., Kamendulis, L.M., . . . Saykin, A.J. (2012). Alterations in brain structure and function in breast cancer survivors: Effect of post-chemotherapy interval and relation to oxidative DNA damage. *Breast Cancer Research and Treatment, 137*, 493–502. doi:10.1007/s10549-012-2385-x

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan (D-KEFS) executive function system, examiners manual*. San Antonio, TX: Psychological Corporation.

DNA Genotek. (2012). Laboratory protocol for manual purification of DNA from whole sample. Retrieved from http://www.dnagenotek.com/US/support/protocols-oragene-discover.html

Downie, F.P., Mar Fan, H.G., Houédé-Tchen, N., Yi, Q., & Tannock, I.F. (2006). Cognitive function, fatigue, and menopausal symptoms in breast cancer patients receiving adjuvant chemotherapy: Evaluation with patient interview after formal assessment. *Psycho-Oncology, 15*, 921–930. doi:10.1002/pon.1035

Edge, S.B., Byrd, D.R., Compton, C.C., Fritz, A.G., Greene, F.L., & Trotti, A. (Eds.). (2010). *AJCC cancer staging manual* (7th ed.). New York, NY: Springer.

Falleti, M.G., Sanfilippo, A., Maruff, P., Weih, L., & Phillips, K.A. (2005). The nature and severity of cognitive impairment associated with adjuvant chemotherapy in women with breast cancer: A meta-analysis of the current literature. *Brain and Cognition, 59*, 60–70. doi:10.1016/j.bandc.2005.05.001

Farrer, L., Cupples, L., Haines, J., Hyman, B., Kukull, W., Mayeux, R., . . . Van Duijn, C. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. *JAMA, 278*, 1349–1356. doi:10.1001/jama.1997.03550160069041

Finkel, T., & Holbrook, N.J. (2000). Oxidants, oxidative stress and the biology of ageing. *Nature, 408*, 239–247. doi:10.1038/35041687

Flory, J.D., Manuck, S.B., Ferrell, R.E., Ryan, C.M., & Muldoon, M.F. (2000). Memory performance and the apolipoprotein E polymorphism in a community sample of middle-aged adults. *American Journal of Medical Genetics, 96*, 707–711. doi:10.1002/1096-8628(20001204)96:6<707

Hamed, E.A., Zakhary, M.M., & Maximous, D.W. (2012). Apoptosis, angiogenesis, inflammation, and oxidative stress: Basic interactions in patients with early and metastatic breast cancer. *Journal of Cancer Research and Clinical Oncology, 138*, 999–1009. doi:10.1007/s00432-012-1176-4

Han, S.D., & Bondi, M.W. (2008). Revision of the apolipoprotein E compensatory mechanism recruitment hypothesis. *Alzheimer's and Dementia, 4*, 251–254. doi:10.1016/j.jalz.2008.02.006

Hayek, T., Oiknine, J., Brook, J.G., & Aviram, M. (1994). Increased plasma and lipoprotein lipid peroxidation in Apo E-deficient mice. *Biochemical and Biophysical Research Communications, 201*, 1567–1574. doi:10.1006/bbrc.1994.1883

Hermelink, K., Untch, M., Lux, M.P., Kreienberg, R., Beck, T., Bauerfeind, I., & Münzel, K. (2007). Cognitive function during neoadjuvant chemotherapy for breast cancer: Results of a prospective, multicenter, longitudinal study. *Cancer, 109*, 1905–1913. doi:10.1002/cncr.22610

Hofer, S.M., Christensen, H., Mackinnon, A.J., Alisa, E., Jorm, A.F., Henderson, A.S., & Easteal, S. (2002). Change in cognitive functioning associated with ApoE genotype in a community sample of older adults. *Psychology and Aging, 17*, 194–208. doi:10.1037/0882-7974.17.2.194

Howlader, N., Noone, A.M., Krapcho, M., Neyman, N., Aminou, R., Altekruse, S.F., . . . Cronin, K.A. (2011). *SEER cancer statistics review, 1975-2008 (Vintage 2009 Populations)*. Retrieved from http://seer.cancer.gov/csr/1975_2008/

Hubacek, J.A., Pitha, J., Skodova, Z., Adamkova, V., Lanska, V., & Poledne, R. (2001). A possible role of apolipoprotein E polymorphism in predisposition to higher education. *Neuropsychobiology, 43*, 200–203. doi:10.1159/000054890

Hurria, A., Goldfarb, S., Rosen, C., Holland, J., Zuckerman, E., Lachs, M.S., . . . Hudis, C. (2006). Effect of adjuvant breast cancer chemotherapy on cognitive function from the older patient's perspective. *Breast Cancer Research and Treatment, 98*, 343–348. doi:10.1007/s10549-006-9171-6

Izaks, G.J., Gansevoort, R.T., Van der Knaap, A.M., Navis, G., Dullaart, R.P.F., & Slaets, J.P.J. (2011). The association of APOE genotype with cognitive function in persons aged 35 years or older. *PLOS One, 6*, e27415. doi:10.1371/journal.pone.0027415

Janelsins, M.C., Kohli, S., Mohile, S.G., Usuki, K., Ahles, T., & Morrow, G.R. (2012). An update on cancer- and chemotherapy-related cognitive dysfunction. *Seminars in Oncology, 38*, 431–438. doi:10.1053/j.seminoncol.2011.03.014.AN

Jenkins, V., Shilling, V., Deutsch, G., Bloomfield, D., Morris, R., Allan, S., . . . Winstanley, J. (2006). A 3-year prospective study of the effects of adjuvant treatments on cognition in women with early stage breast cancer. *British Journal of Cancer, 94*, 828–834. doi:10.1038/sj.bjc.6603029

Jofre-Monseny, L., Minihane, A.M., & Rimbach, G. (2008). Impact of apoE genotype on oxidative stress, inflammation and disease risk. *Molecular Nutrition and Food Research, 52*, 131–145. doi:10.1002/mnfr.200700322

Jolivalt, C., Leininger-Muller, B., Bertrand, P., Herber, R., Christen, Y., & Siest, G. (2000). Differential oxidation of apolipoprotein E isofroms and interaction with phospholipids. *Free Radical Biology and Medicine, 28*, 129–140. doi:10.1016/S0891-5849(99)00232-4

Jorm, A.F., Mather, K.A., Butterwork, P., Anstey, K.J., Christensen, H., & Easteal, S. (2007). APOE genotype and cognitive functioning in a large age-stratified population sample. *Neuropsychology, 21*, 1–8. doi:10.1037/0894-4105.21.1.1

Joshi, G., Sultana, R., Tangpong, J., Cole, M.P., St. Clair, D.K., Vore, M., . . . Butterfield, D.A. (2005). Free radical mediated oxidative stress and toxic side effects in brain induced by the anti cancer drug adriamycin: Insight into chemobrain. *Free Radical Research, 39*, 1147–1154. doi:10.1080/10715760500143478

Kasapovic, J., Pejic, S., Stojiljkovic, V., Todorovic, A., Radoševic-Jelic, L., Saicic, Z.S., & Pajovic, S.B. (2010). Antioxidant status and lipid peroxidation in the blood of breast cancer patients of different ages after chemotherapy with 5-fluorouracil, doxorubicin and cyclophosphamide. *Clinical Biochemistry, 43*(16–17), 1287–1293. doi:10.1016/j.clinbiochem.2010.08.009

Lezak, M.D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.

Mahley, R.W., Innerarity, T.L., Rall, S.C., & Weisgraber, K.H. (1984). Plasma lipoproteins: Apolipoprotein structure and function. *Journal of Lipid Research, 25*, 1277–1294. Retrieved from http://www.jlr.org/content/25/12/1277.long

Mariani, E., Polidori, M.C., Cherubini, A., & Mecocci, P. (2005). Oxidative stress in brain aging, neurodegenerative and vascular diseases: An overview. *Journal of Chromatography, 827*, 65–75. doi:10.1016/j.jchromb.2005.04.023

Matthews, C.G. (1964). *Adult neuropsychological test battery*. Madison, WI: University of Wisconsin Medical Center.

McNair, D., Lorr, M., & Droppleman, L.F. (1992). *EdITS manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service.

Mehnert, A., Scherwath, A., Schirmer, L., Schleimer, B., Petersen, C., Schulz-Kindermann, F., . . . Koch, U. (2007). The association between neuropsychological impairment, self-perceived cognitive deficits, fatigue and health related quality of life in breast cancer survivors following standard adjuvant versus high-dose chemotherapy. *Patient Education and Counseling, 66,* 108–118. doi:10.1016/j.pec.2006.11.005

Miller, S.A., Dykes, D.D., & Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research, 16,* 1215. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC334765/

Miyata, M., & Smith, J.D. (1996). Apolipoprotein E allele-specific antioxidant activity and effects on cytotoxicity by oxidation insults and β-amyloid peptides. *Nature Genetics, 14,* 55–61. doi:10.1038/ng0996-55

Mondadori, C.R.A., De Quervain, D.J.F., Buchmann, A., Mustovic, H., Wollmer, M.A., Schmidt, C.F., . . . Henke, K. (2007). Better memory and neural efficiency in young apolipoprotein E epsilon4 carriers. *Cerebral Cortex, 17,* 1934–1947. doi:10.1093/cercor/bhl103

Munir, F., Burrows, J., Yarker, J., Kalawsky, K., & Bains, M. (2010). Women's perceptions of chemotherapy-induced cognitive side affects on work ability: A focus group study. *Journal of Clinical Nursing, 19*(9–10), 1362–1370. doi:10.1111/j.1365-2702.2009.03006.x

Myers, J.S. (2012). Chemotherapy-related cognitive impairment: The breast cancer experience [Online exclusive]. *Oncology Nursing Forum, 39,* E31–E40. doi:10.1188/12.ONF.E31-E40

Nelson, H. (1981). *Nelson Adult Reading Test (NART) manual*. Windsor, Ontario: NFER-Nelson.

Nilsson, L.G., Nyberg, L., & Bäckman, L. (2002). Genetic variation in memory functioning. *Neuroscience and Biobehavioral Reviews, 26,* 841–848. doi:10.1016/S0149-7634(02)00070-2

Osterrieth, P.A. (1944). Le test de copie d'une figure complexe. *Archives de Psychologie, 30,* 206–356.

Owen, A.M., Sahakian, B.J., Semple, J., Polkey, C.E., & Robbins, T.W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia, 33,* 1–24. doi:10.1016/0028-3932(94)00098-A

Pedersen, W.A., Chan, S.L., & Mattson, M.P. (2000). A mechanism for the neuroprotective effect of apolipoprotein E: Isoform-specific modification by the lipid peroxidation product 4-hydroxynonenal. *Journal of Neurochemistry, 74,* 1426–1433. doi:10.1046/j.1471-4159.2000.0741426.x

Reitan, R., & Wolfson, D. (1985). *The Halstead-Reitan neuropsychological test battery*. Tucson, AZ: Neuropsychology Press.

Rey, A. (1964). L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives in Psychology, 1252,* 340–382.

Richard, F., & Amouyel, P. (2001). Genetic susceptibility factors for Alzheimer's disease. *European Journal of Pharmacology, 412*(1), 1–12. doi:10.1016/S0014-2999(00)00903-1

Sadigh-Eteghad, S., Talebi, M., & Farhoudi, M. (2012). Association of apolipoprotein E epsilon 4 allele with sporadic late onset Alzheimer`s disease. A meta-analysis. *Neurosciences, 17,* 321–326.

Schagen, S.B., Van Dam, F.S., Muller, M.J., Boogerd, W., Lindeboom, J., & Bruning, P.F. (1999). Cognitive deficits after postoperative adjuvant chemotherapy for breast carcinoma. *Cancer, 85,* 640–650. doi: 10.1002/(SICI)1097-0142(19990201)85:3<640::AID-CNCR14>3.0.CO;2-G

Schilder, C.M., Eggens, P.C., Seynaeve, C., Linn, S.C., Boogerd, W., Gundy, C.M., . . . Schagen, S.B. (2009). Neuropsychological functioning in postmenopausal breast cancer patients treated with tamoxifen or exemestane after AC-chemotherapy: Cross-sectional findings from the neuropsychological TEAM-side study. *Acta Oncologica, 48,* 76–85. doi:10.1080/02841860802314738

Schilder, C.M., Seynaeve, C., Beex, L.V., Boogerd, W., Linn, S.C., Gundy, C.M., . . . Schagen, S.B. (2010). Effects of tamoxifen and exemestane on cognitive functioning of postmenopausal patients with breast cancer: Results from the neuropsychological side study of the tamoxifen and exemestane adjuvant multinational trial. *Journal of Clinical Oncology, 28,* 1294–1300. doi:10.1200/JCO.2008.21.3553

Shilling, V., & Jenkins, V. (2007). Self-reported cognitive problems in women receiving adjuvant therapy for breast cancer. *European Journal of Oncology Nursing, 11,* 6–15. doi:10.1016/j.ejon.2006.02.005

Small, B.J., Rawson, K.S., Walsh, E., Jim, H.S.L., Hughes, T.F., Iser, L., . . . Jacobsen, P.B. (2011). Catechol-O-methyltransferase genotype modulates cancer treatment-related cognitive deficits in breast cancer survivors. *Cancer, 117,* 1369–1376. doi:10.1002/cncr.25685

Stilley, C.S., Bender, C.M., Dunbar-Jacob, J., Sereika, S., & Ryan, C.M. (2011). The impact of cognitive function on medication management: Three studies. *Health Psychology, 29,* 50–55. doi:10.1037/a0016940

Strehlow, K., Rotter, S., Wassmann, S., Adam, O., Grohé, C., Laufs, K., . . . Nickenig, G. (2003). Modulation of antioxidant enzyme expression and function by estrogen. *Circulation Research, 93,* 170–177. doi:10.1161/01.RES.0000082334.17947.11

Tchen, N., Juffs, H.G., Downie, F.P., Yi, Q.L., Hu, H., Chemerynsky, I., . . . Tannock, I.F. (2003). Cognitive function, fatigue, and menopausal symptoms in women receiving adjuvant chemotherapy for breast cancer. *Journal of Clinical Oncology, 21,* 4175–4183. doi:10.1200/JCO.2003.01.119

Unfer, T.C., Conterato, G.M., Da Silva, J.C., Duarte, M.M., & Emanuelli, T. (2006). Influence of hormone replacement therapy on blood antioxidant enzymes in menopausal women. *Clinica Chimica Acta, 369,* 73–77. doi:10.1016/j.cca.2006.01.006

Vardy, J., Wefel, J.S., Ahles, T., Tannock, I.F., & Schagen, S.B. (2008). Cancer and cancer-therapy related cognitive dysfunction: An international perspective from the Venice cognitive workshop. *Annals of Oncology, 19,* 623–629. doi:10.1093/annonc/mdm500

Von Ah, D., Habermann, B., Carpenter, J.S., & Schneider, B.L. (2013). Impact of perceived cognitive impairment in breast cancer survivors. *European Journal of Oncology Nursing, 17,* 236–241. doi:10.1016/j.ejon.2012.06.002

Walker, C.H., Drew, B.A., Antoon, J.W., Kalueff, A.V., & Beckman, B.S. (2012). Neurocognitive effects of chemotherapy and endocrine therapies in the treatment of breast cancer: Recent perspectives. *Cancer Investigation, 30,* 135–148. doi:10.3109/07357907.2011.636116

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York, NY: The Psychological Corporation.

Wefel, J.S., Lenzi, R., Theriault, R., Buzdar, A.U., Cruickshank, S., & Meyers, C.A. (2004). "Chemobrain" in breast carcinoma?: A prologue. *Cancer, 101,* 466–475. doi:10.1002/cncr.20393

Wefel, J.S., Saleeba, A.K., Buzdar, A.U., & Meyers, C.A. (2010). Acute and late onset cognitive dysfunction associated with chemotherapy in women with breast cancer. *Cancer, 116,* 3348–3356. doi:10.1002/cncr.25098

Wehling, E., Lundervold, A.J., Standnes, B., Gjerstad, L., & Reinvang, I. (2007). APOE status and its association to learning and memory performance in middle aged and older Norwegians seeking assessment for memory deficits. *Behavioral and Brain Functions, 3,* 57. doi:10.1186/1744-9081-3-57

Wilson, B., Cockburn, J., Baddeley, A., & Hiorns, R. (1989). The development and validation of a test battery for detecting and monitoring everyday memory problems. *Journal of Clinical and Experimental Neuropsychology, 11,* 855–870. doi:10.1080/01688638908400940

Wishart, H.A., Saykin, A.J., Rabin, L.A., Santulli, R.B., Flashman, L.A., Guerin, S.J., . . . Mcallister, T.W. (2006). Increased brain activation during working memory in cognitively intact adults with the APOE ε4 allele. *American Journal of Psychiatry, 163,* 1603–1610. doi:10.1176/appi.ajp.163.9.1603

Yu, Y.W., Lin, C.H., Chen, S.P., Hong, C.J., & Tsai, S.J. (2000). Intelligence and event-related potentials for young female human volunteer apolipoprotein E epsilon4 and non-epsilon4 carriers. *Neuroscience Letters, 294,* 179–181.

APPENDIX G

LICENSE AGREEMENT FOR MANUSCRIPT #2

RE: ONF Reuse Request

Mike Minjock <mminjock@ons.org>

Thu 3/31/2016 6:50 PM

To:Koleck, Theresa Ann <tat30@pitt.edu>;

Dear Theresa,

Yes. You may include a copy of the article in your dissertation document.

Sincerely,

Mike Minjock
Licensing Manager
Oncology Nursing Society
125 Enterprise Drive
Pittsburgh, PA 15275-1214
+1-412-859-6251 (phone)
+1-412-859-6163 (fax)
mminjock@ons.org
www.ons.org

From: Koleck, Theresa Ann [mailto:tat30@pitt.edu]
Sent: Tuesday, March 29, 2016 10:18 AM
To: pubpermissions <pubpermissions@ons.org>
Subject: ONF Reuse Request

Dear ONS,

I am writing to request permission to include a copy of an article, Apolipoprotein E Genotype and Cognitive Function in Postmenopausal Women With Early-Stage Breast Cancer, which was published in Oncology Nursing Forum in 2014 during my PhD studies, in my electronic dissertation document. I am the first author on this publication.

Please find the completed request form attached.

Thank you for your assistance.

Sincerely,
Theresa

Theresa A. Timcheck Koleck, PhD(c), BSN, RN
Predoctoral Scholar
F31NR014590
School of Nursing
University of Pittsburgh
(412) 862-7144
tat30@pitt.edu

Oncology Nursing Society

125 Enterprise Drive • Pittsburgh, PA 15275-1214
Toll Free: 866-257-4ONS • Phone: 412-859-6100 • Fax: 412-859-6163
pubpermissions@ons.org • www.ons.org

Request to Reuse Oncology Nursing Society (ONS) Content

Permission from ONS is required to photocopy, distribute, post online, or otherwise share ONS copyrighted content or to reuse articles, tables, figures, excerpts, and other content from ONS books, journals, websites, web courses, and other ONS products. To request permission, please complete the form below and email it as an attachment to pubpermissions@ons.org. **A royalty or licensing fee may apply. Please allow 2–4 weeks for processing.**

DATE OF REQUEST: March 29, 2016

IDENTIFICATION OF INFORMATION TO BE REUSED:

Title of publication (book or journal), including volume and issue numbers: Oncology Nursing Forum, 41(6)
Title of book chapter or article: Apolipoprotein E Genotype and Cognitive Function in Postmenopausal Women With Early-Stage Breast Cancer
Author(s)/editor(s): Koleck TA, Bender CM, Sereika SM, Ahrendt G, Jankowitz RC, McGuire KP, Ryan CM, Conley YP.
Year of publication: 2014
Page number(s): E313-25
What item(s) do you want to reproduce? (list table/figure number, appendix, entire article, etc.) entire article

INFORMATION ABOUT REQUESTOR:

Name: Theresa A. Koleck
Company or institution: University of Pittsburgh, School of Nursing
Street address: 440 Victoria Building, 3500 Victoria Street
City: Pittsburgh **State:** PA **Zip/postal code:** 15261 **Country:** USA
Phone number: 412-862-7144
E-mail address: tat30@pitt.edu

Are you (or the person you are representing) a contributing author of the ONS material you wish to reuse? Yes

USE/REPUBLICATION INFORMATION:

Material will be [check one]

☐ Reprinted for use in another publication
☐ Adapted for use in another publication [attach copy of adaptation]
☐ Photocopied for distribution [number of copies]
☐ Posted online: If yes, give time duration, name of website, and web address
☒ Other [please explain] reuse in Theresa Koleck's dissertation document

Company or institution that will publish or otherwise reuse the content: University of Pittsburgh
Purpose (staff training, classroom teaching, patient education, etc.): dissertation document
Estimated number of potential viewers: 5
Who will the viewers be? Individual searching dissertations
Will a purchase price, tuition, registration cost, or other fee be charged to view?: No
Expected date of publication or use: April 25, 2017
Number of individual hospitals in your system, facilities, schools, or other institutions that will use: 0
If hospital(s) or other care facilities, how many "patient beds" at each location?: N/A
Distribution (North America only or worldwide): worldwide
Language(s): English

If a publication, type of publication in which material will be reproduced (book, journal, website, etc.): dissertation

Title: Cognitive Function and Breast Cancer: Genomics and Disease Characteristics
Author: Theresa A. Koleck
Publisher: University of Pittsburgh
Expected print run: None
Will electronic or online version be available?: Yes

- **If yes, please provide the estimated number of online viewers:** 5

Web address if "online only" or if online version of print version will be available: electronic thesis and dissertation

Further comments or explanation so we fully understand your request: This first author publication was written and published during my PhD studies and I would like to include a copy of it in my electronic dissertation document for the University of Pittsburgh.

If granted, permission would be subject to the following conditions:

1. Reproduced material would need to be clearly identified with the name of the copyright holder, author(s)/editor(s), year of publication, book/chapter/article title, and page number(s) where material appeared.
2. Permission would be granted for nonexclusive, one-time use only, as described in the request. Electronic or any other usage would require a separate request.
3. Permission would be granted for use in the English language only. Translations of materials would require a separate request.
4. Permission could only be granted if the material is original to an Oncology Nursing Society publication. If the requested material has appeared in our publication with credit or acknowledgment to another source, it would be the requestor's responsibility to secure permission from that source, as well.

APPENDIX H

MANUSCRIPT #3: IDENTIFICATION AND PRIORITIZATION OF CANDIDATE GENES FOR SYMPTOM VARIABILITY IN BREAST CANCER SURVIVORS BASED ON DISEASE CHARACTERISTICS AT THE CELLULAR LEVEL

Breast Cancer: Targets and Therapy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

METHODOLOGY

Identification and prioritization of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level

This article was published in the following Dove Press journal:
Breast Cancer: Targets and Therapy
8 March 2016
Number of times this article has been viewed

Theresa A Koleck[1]
Yvette P Conley[2]

[1]School of Nursing, [2]Department of Human Genetics, School of Nursing and Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

Abstract: Research is beginning to suggest that the presence and/or severity of symptoms reported by breast cancer survivors may be associated with disease-related factors of cancer. In this article, we present a novel approach to the identification and prioritization of biologically plausible candidate genes to investigate relationships between genomic variation and symptom variability in breast cancer survivors. Cognitive dysfunction is utilized as a representative breast cancer survivor symptom to elucidate the conceptualization of and justification for our cellular, disease-based approach to address symptom variability in cancer survivors. Initial candidate gene identification was based on genes evaluated as part of multigene expression profiles for breast cancer, which are commonly used in the clinical setting to characterize the biology of cancer cells for the purpose of describing overall tumor aggressiveness, prognostication, and individualization of therapy. A list of genes evaluated within five multigene expression profiles for breast cancer was compiled. In order to prioritize candidate genes for investigation, genes used in each profile were compared for duplication. Twenty-one genes (*BAG1*, *BCL2*, *BIRC5*, *CCNB1*, *CENPA*, *CMC2*, *DIAPH3*, *ERBB2*, *ESR1*, *GRB7*, *MELK*, *MKI67*, *MMP11*, *MYBL2*, *NDC80*, *ORC6*, *PGR*, *RACGAP1*, *RFC4*, *RRM2*, and *SCUBE2*) are utilized in two or more profiles, including five genes (*CCNB1*, *CENPA*, *MELK*, *MYBL2*, and *ORC6*) used in three profiles. To ensure that the parsimonious 21 gene set is representative of the more global biological hallmarks of cancer, an Ingenuity Pathway Analysis was conducted. Evaluation of genes known to impact pathways involved with cancer development and progression provide a means to evaluate the overlap between the biological underpinnings of cancer and symptom development within the context of cancer.

Keywords: breast neoplasms, biological markers, genes, signs and symptoms, cognition

Introduction

Many cancer survivors experience a variety of disruptive and burdensome symptoms, including fatigue, pain, altered sleep, mood dysregulation, and cognitive dysfunction long into survivorship.[1–3] Although our ability to describe the duration, frequency, and severity of symptoms related to cancer and cancer treatments has vastly improved, our understanding of the mechanisms that influence symptom variability and our ability to personalize symptom prediction for an individual cancer survivor and intervene effectively remain limited. While still in its infancy, research is beginning to suggest that the presence and/or severity of symptoms reported by cancer survivors may not be solely the result of cancer treatments, but may be associated with disease-related

Correspondence: Theresa A Koleck
Department of Health Promotion and Development, School of Nursing, University of Pittsburgh, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA
Tel +1 412 383 7641
Fax +1 412 624 8521
Email tat30@pitt.edu

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/BCTT.S88434

factors of cancer and/or host characteristics that predispose an individual to cancer as well as a particular symptom.[4–6]

Breast cancer survivors have been the focus of a large proportion of cancer survivor symptom research. Studies conducted with breast cancer survivors on the symptom of cognitive dysfunction especially, are contributing an increasing amount of evidence in support of our hypothesis that disease-related factors of cancer and/or host characteristics that influence cancer development or progression contribute to the presence and severity of symptoms experienced by cancer survivors.

Cognitive dysfunction research in breast cancer survivors has traditionally concentrated on the direct neurotoxic effect of chemotherapeutic agents on the brain. Often referred to as "chemo brain" or "chemo fog", short- and long-term cognitive changes have been well-documented in women with breast cancer receiving chemotherapy.[7] With the increasing use of antiestrogen therapies for prevention of breast cancer recurrence, the influence of estrogen and the use of antiestrogen therapies on cognitive decline in breast cancer survivors have become an additional focus of research on treatment-related cognitive changes.[8–14]

However, more recent research demonstrates that cognitive changes may actually occur in breast cancer survivors prior to initiation of adjuvant chemotherapy and/or antiestrogen therapy. In a study of 109 women with breast cancer scheduled to receive chemotherapy, Hermelink et al[15] found that group mean scores were significantly poorer than test norms on five out of twelve cognitive tests before the start of treatment. In addition, 33 survivors scored in the lower fifth percentile of test norms on two or more cognitive tests unrelated to depression, anxiety, or self-reported cognitive problems. Similarly, Wefel et al[16] reported that 29 out of 84 breast cancer survivors diagnosed with stage 1–3a breast cancer were classified as cognitively impaired (ie, multiple cognitive tests with z-scores ≤ -1.5 or a single test with z-score ≤ -2.0) before receiving chemotherapy compared to normative data.

Even more compelling are findings from studies comparing the cognitive function of breast cancer survivors to healthy controls. Ahles et al[17] found that women with invasive breast cancer had poorer overall cognitive performance compared to women with noninvasive breast cancer and healthy controls. Bender et al[18] also reported pretreatment differences in cognitive function in the domains of verbal learning and memory and attention between women with breast cancer prescribed antiestrogen therapy with or without chemotherapy and healthy controls matched on age and years of education. Although not statistically significant, Schagen et al[19] found that 16.4% of survivors prescribed chemotherapy and 29.8% of survivors with stage 1 breast cancer who were prescribed no systemic treatment displayed cognitive impairment before initiation of adjuvant treatment compared to 10% of healthy controls.

While multiple factors potentially predict cognitive function in women with breast cancer prior to adjuvant therapy, we hypothesize that these pretreatment findings suggest that disease-related factors inherent in breast cancer and/or host characteristics that predispose an individual to cancer as well as cognitive dysfunction may be a major determinant of cognitive changes in women with breast cancer. Additionally, only a subset of breast cancer survivors appears to be affected by cognitive dysfunction. We further hypothesize that heterogeneity in the biology of breast cancers at the cellular level could account for a significant proportion of reported discrepancies in cognitive function among survivors.

One common clinical tool used to evaluate the underlying biology of breast cancer cells is the prognostic multigene expression profile for breast cancer. These profiles enhance knowledge received from traditional tumor features and utilize predictive algorithms of tumor gene expression levels to individualize treatment through estimation of adjuvant therapy benefit and distant cancer recurrence risk. Thus, each prognostic multigene expression profile is comprised of genes that play an important role in breast cancer aggressiveness and progression, and, consequently, represent ideal candidates for a genetic association study exploring our hypotheses (Figure 1).

In this article, we specifically aim to present a novel approach, based on genes examined in prognostic multigene expression profiles for breast cancer, to the identification and prioritization of biologically plausible candidate genes for future investigations of the association between genetic variation and

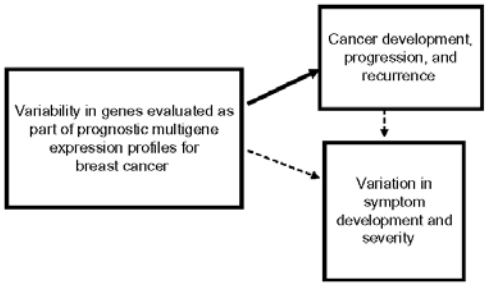


Figure 1 Conceptual model of using variability in genes evaluated as part of prognostic multigene expression profiles for breast cancer to test the hypothesis that heterogeneity in the biology of breast cancers at the cellular level could account for symptom variation.
Note: Dashed arrows represent relationships to be tested in future investigations.

Table 1 Prognostic multigene expression profiles

Profile	Provider	Number of cancer genes evaluated	Clinical utility	Patient eligibility	Website	Reference
11-gene expression signature	Breast Cancer Index[SM] bioTheranostics, Inc. (San Diego, CA, USA)	7 (5 gene molecular grade index; 2 gene HOXB13: IL17BR index, an expression ratio biomarker)	Predicts late (5–10 year) and overall (0–10 year) recurrence risk as well as likelihood of endocrine therapy benefit; stratifies patients into high or low risk of distant recurrence and high or low likelihood of benefit from extended endocrine therapy	Patients with ER (+), lymph node (−/+, 1–3 positive nodes), early stage breast cancer who are distant recurrence-free	http://breastcancerindex.com	20–22
14-gene prognostic expression signature	Not currently supported	14	Predicts distant metastasis; guides treatment decisions related to adjuvant therapy	Patients with ER (+), lymph node (−), early stage breast cancer	NA	23
21-gene breast cancer assay	Oncotype DX® Genomic Health®, Inc. (Redwood City, CA, USA)	16 (for invasive breast cancer) 7 (for DCIS)	Predicts 10-year risk of distant recurrence and likelihood of chemotherapy benefit for invasive breast cancer; gene expression levels are aggregated and provided as a scaled Recurrence Score® result (on a 0–100 scale); patients are stratified into high, intermediate, or low risk groups; also provides quantitative single gene scores for ER, PR, and HER2; predicts 10-year risk of distant recurrence (DCIS or invasive) for DCIS	Pre- or post-menopausal patients with ER (+), lymph node (−), HER2 (−), stage I, II or IIIa invasive breast cancer Post-menopausal patients with lymph node (+, 1–3 positive nodes), HR (+), HER2 (−), stage I, II or IIIa invasive breast cancer DCIS treated by lumpectomy, not mastectomy, regardless of Tamoxifen status (+/−)	http://breast-cancer.oncotypedx.com	24,25
50-gene breast cancer prognostic gene signature assay	Prosigna® NanoString® Technologies, Inc. (Seattle, WA, USA)	50	Predicts 10-year distant recurrence-free survival using gene-algorithm generated Prosigna® Score (0–100); stratifies patients into high, intermediate, or low risk groups. Provides risk group classification to facilitate interpretation of Prosigna® score with clinical outcomes	Post-menopausal patients with HR (+), lymph node (−), stage I or II invasive breast cancer to be treated with adjuvant endocrine therapy alone Post-menopausal patients with HR (+), lymph node (+, 1–3 positive nodes), stage II invasive breast cancer treated with adjuvant endocrine therapy alone	http://www.prosigna.com	26
70-gene breast cancer recurrence assay	MammaPrint® Agendia® (Irvine, CA, USA)	70	Predicts 5- and 10-year distant recurrence and guides treatment decisions, including potential chemotherapy benefit; stratifies patients into low or high risk groups	Patients with ER (+/−), lymph node (−), stage I or II invasive breast cancer tumor ≤5 cm in size	http://www.agendia.com/healthcare-professionals/breast-cancer/mammaprint	27

Notes: Information regarding Prosigna® is provided courtesy of NanoString Technologies, Inc., www.Prosigna.com. © 2014–15 NanoString Technologies, Inc. All rights reserved. Information regarding Oncotype DX® is provided courtesy of Genomic Health, Inc., www.genomichealth.com. © 2015 Genomic Health, Inc., All rights reserved.
Abbreviations: DCIS, ductal carcinoma in situ; ER, estrogen receptor; HR, hormone receptor; PR, progesterone receptor.

symptoms experienced by breast cancer survivors. We will focus on 1) characteristics of five different prognostic multigene expression profiles for breast cancer, 2) prioritization of candidate genes replicated in two or more profiles, 3) biological functions of our identified, parsimonious candidate gene set, and 4) a discussion of the potential expanded clinical utility of prognostic multigene expression profiles for breast cancer, and, more generally, cancer symptom prediction.

Methods

Selected breast-cancer-related prognostic multigene expression profiles

Prognostic multigene expression profiles use tumor gene expression levels to evaluate the underlying biology of breast cancer cells and predict long-term outcomes and potential benefit of additional adjuvant therapy. Several groups have developed prognostic multigene expression profiles for breast cancer: the eleven-gene expression signature (marketed as the Breast Cancer Index℠ by bioTheranostics, Inc., San Diego, CA, USA),[28] the 14-gene prognostic expression signature (described in Tutt et al[23]), the 21-gene breast cancer assay (marketed as the Oncotype DX® Breast Cancer Assay by Genomic Health®, Inc., Redwood City, CA, USA),[29,30] the 50-gene breast cancer prognostic gene signature assay (marketed as the Prosigna® Breast Cancer Prognostic Gene Signature Assay by NanoString® Technologies, Inc., Seattle, WA, USA) based on the PAM50 Breast Cancer Intrinsic Classifier,[31] and the 70-gene breast cancer recurrence assay (marketed as the MammaPrint® 70 Gene Breast Cancer Recurrence Assay by Agendia®, Irvine, CA, USA).[32,33]

The number of cancer genes utilized in each profile varies greatly, ranging from 7 in the eleven-gene expression signature to 70 in the 70-gene breast cancer recurrence assay. All of the prognostic multigene expression profiles provide predictions of 5- and/or 10-year distant breast cancer recurrence risk, except the 14-gene prognostic expression signature, which is specifically intended for prediction of distant metastasis. With the exception of the 50-gene breast cancer prognostic gene

Table 2 Genes utilized in two or more prognostic multigene expression profiles as indicated by X

Gene	11-gene expression profile	14-gene prognostic expression signature	21-gene breast cancer assay	50-gene breast cancer prognostic gene signature assay	70-gene breast cancer recurrence assay	Gene function
BAG1			X	X		Enhances antiapoptotic effect of BCL2
BCL2			X	X		Blocks the apoptotic death of certain cells
BIRC5			X	X		Encodes regulatory proteins that prevent apoptosis
CCNB1[a]		X	X	X		Encodes a regulatory protein involved in mitosis
CENPA[a]	X	X			X	Encodes for a centromere protein; histone H3 variant
CMC2		X			X	Potential involvement in mitochondrial cytochrome c oxidase biogenesis[34]
DIAPH3		X			X	Involved in actin remodeling and regulation of cell movement and adhesion
ERBB2			X	X		Encodes HER2, an epidermal growth factor receptor protein
ESR1			X	X		Encodes an estrogen receptor
GRB7			X	X		Encodes a growth factor receptor-binding protein
MELK[a]		X		X	X	Involved in cell cycle regulation, apoptosis, and splicing regulation[35–37]
MKI67			X	X		Involved in cellular proliferation
MMP11			X	X		Involved in extracellular matrix breakdown
MYBL2[a]		X	X	X		Encodes a nuclear protein; involved in cell cycle progression
NDC80				X	X	Organization and stabilization of microtubule–kinetochore attachments
ORC6[a]		X		X	X	Involved in chromosome replication and segregation
PGR			X	X		Encodes a progesterone receptor; mediates effects of progesterone
RACGAP1	X	X				Involved in cytokinesis initiation and control of cellular growth
RFC4		X			X	Required for elongation of primed DNA templates
RRM2	X		X			Catalyzes the formation of deoxyribonucleotides from ribonucleotides
SCUBE2			X		X	Potential breast tumor suppressor gene[30]

Notes: Information on gene function was obtained from the NCBI's Gene Database[39] unless noted otherwise. [a]Indicates a gene used in three expression profiles.
Abbreviation: NCBI, National Center for Biotechnology Information.

signature assay, which is not purposed to assist in the selection of optimal therapy, results from the remaining prognostic multigene expression profiles are intended to guide clinical treatment decisions, relaying the benefit of additional adjuvant chemotherapy and/or antiestrogen therapy. Table 1 compares important characteristics of the five prognostic multigene expression profiles, including number of genes evaluated, clinical utility, and patient eligibility.

Identification and prioritization of candidate genes

A list of genes evaluated within each of the five prognostic multigene expression profiles was compiled. Lists of genes were obtained from the following locations in March 2014: the eleven-gene expression profile (Jerevall et al),[28] the 14-gene prognostic expression signature (Tutt et al),[23] the 21-gene breast cancer assay (http://breast-cancer.oncotypedx.com),

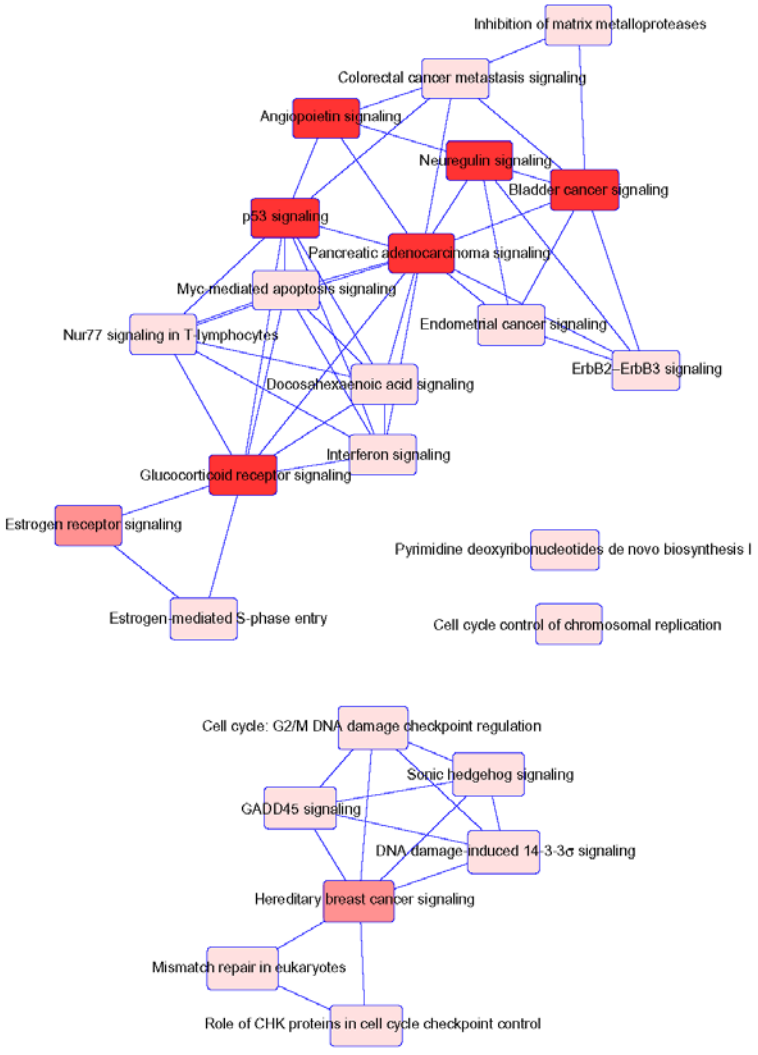


Figure 2 Overlapping canonical pathways map representing shared biology among the identified candidate genes.
Notes: Connected canonical pathways share one or more genes in common. The brighter the red of the node, the more significant the canonical pathway in the gene set. The canonical pathways map was generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity).

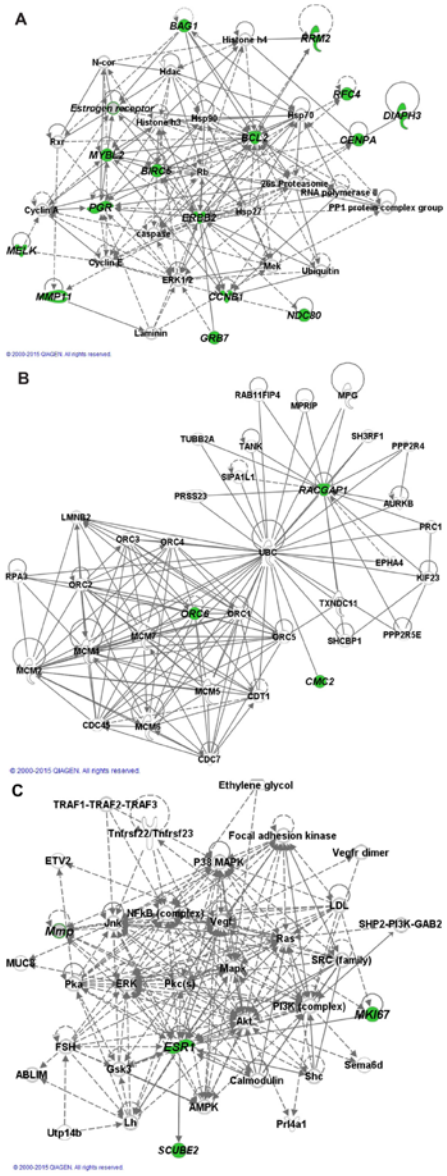


Figure 3 Gene–gene networks generated by pathway analysis.
Notes: The networks were generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). Identified candidate genes are highlighted in green. All identified candidate genes are included. The main associated functions of each network are as follows: (**A**) cancer, organismal injury and abnormalities, and reproductive system disease; (**B**) DNA replication, recombination, and repair, connective tissue disorders, and dental disease; and (**C**) cellular development, reproductive system development and function, and molecular transport. The node shapes and relationship type legend can be found at http://ingenuity.force.com/ipa/articles/Feature_Description/Legend.

the 50-gene breast cancer prognostic gene signature assay (https://genome.unc.edu/pubsup/breastGEO/), and 70-gene breast cancer recurrence assay (Tian et al).[40] Genes utilized in two or more profiles were prioritized for selection.

Pathway analysis of gene set

In order to ensure adequate representation of the biological capabilities of cancer in the parsimonious high priority candidate gene set, a gene-gene pathway analysis was conducted, using QIAGEN's Ingenuity® Pathway Analysis software (IPA®, QIAGEN Redwood City, CA, USA, www.qiagen.com/ingenuity), to evaluate functional networks. Both direct and indirect relationships were analyzed. All molecules and/or relationships were considered.

Results

Among the five included prognostic gene expression profiles, 127 unique genes were identified. Twenty-one genes (*BAG1*, *BCL2*, *BIRC5*, *CCNB1*, *CENPA*, *CMC2*, *DIAPH3*, *ERBB2*, *ESR1*, *GRB7*, *MELK*, *MKI67*, *MMP11*, *MYBL2*, *NDC80*, *ORC6*, *PGR*, *RACGAP1*, *RFC4*, *RRM2*, and *SCUBE2*) are utilized in two or more of the profiles (Table 2). Five genes (*CCNB1*, *CENPA*, *MELK*, *MYBL2*, and *ORC6*) are used in three profiles. Primary functions of the identified candidate genes are provided in Table 2.

The pathway analysis revealed that the main molecular and cellular functions of the parsimonious, high priority gene set are cell cycle, cellular development, cellular growth and proliferation, cell death and survival, and gene expression. Canonical pathways containing one or more of the identified 21 genes are displayed in Figure 2. Three unique networks were identified (Figure 3). The main associated diseases and functions of the three networks are 1) cancer, organismal injury and abnormalities, and reproductive system disease; 2) DNA replication, recombination, and repair, connective tissue disorders, and dental disease; and 3) cellular development, reproductive system development and function, and molecular transport. The pathway analysis also identified a number of plausible upstream transcription regulators of the identified 21 gene set, including *TP53*, *CDKN1A*, *CDKN2A*, *E2F1*, and *E2F4*.

Discussion and conclusion

In the future, we envision a holistic, personalized health care environment, in which breast cancer survivors receive not only a refined cancer diagnosis and prognosis based on the results of prognostic multigene expression profiles, but genetically tailored preclinical symptom prediction and

proactive symptom management as well. Inspired by findings related to reported changes in cognitive function in women with breast cancer, this project, intended to identify and prioritize biologically plausible candidate genes, represents an initial and integral step in establishing a relationship between genetic variation and symptom variability in breast cancer survivors.

Driven by our hypothesis that symptom variability in breast cancer survivors is related to variation in the biology of cancer cells, we employed the innovative approach detailed in this article to select candidate genes based on prognostic multigene expression profiles for breast cancer. While we believe that all 127 unique genes evaluated as part of the included prognostic multigene expression profiles for breast cancer should be considered as candidates to test the proposed hypothesis, our project revealed considerable overlap in genes represented in the profiles with 21 genes replicated in two or more profiles. Five of the 21 replicated genes are used in three profiles. Because each prognostic multigene expression profile was developed to individualize breast cancer prognostication based on associations between breast cancer tumor gene expression levels and clinically relevant cancer outcomes, including recurrence and metastasis, replication of genes utilized in multiple profiles may be particularly important in describing the heterogeneity of breast cancer tumor cell biology and, consequently, should be prioritized for evaluation.

Nevertheless, by limiting a future investigation to variation in genes replicated in two or more profiles, we risk the inadvertent elimination of one or more of the biologic capabilities that enable tumor growth and metastatic dissemination. Eight biologic hallmarks of cancer have been identified and include resisting cell death, deregulating cellular energetics, sustaining proliferative signaling, evading growth suppressors, avoiding immune destruction, enabling replicative immortality, activating invasion and metastasis, and inducing angiogenesis.[41] To ensure that our parsimonious, high priority gene set broadly incorporated these eight hallmarks, an Ingenuity® Pathway Analysis was conducted. The results of our pathway analysis revealed that the main molecular and cellular functions of the gene set were cell cycle, cellular development, cellular growth and proliferation, cell death and survival, and gene expression. We also found that the majority of the canonical pathways the genes in our identified gene set are contained within are related to cancer/cellular signaling. Furthermore, the genes identified by the analysis as potential upstream regulators of the gene set, *TP53*, *CDKN1A*, *CDKN2A*, *E2F1*, and *E2F4*, all play important roles in cell cycle control and tumor suppression. Due to the overlap between the molecular and cellular functions of the gene set and the hallmarks of cancer, we believe that the biologic hallmarks of cancer are represented in the prioritized gene set.

Interestingly, a further examination of genes, molecules, gene products, and gene complexes/interactions in the identified networks revealed minimal overlap with genes in the prognostic multigene expression profiles. We were surprised to find that only six genes (*CCNE1, CCNE2, FLT1, MCM6, MMP9,* and *PRC1*), beyond those contained in the inputted 21 gene set, are utilized in one of the five included prognostic multigene expression profiles. While unintentional and not the purpose of this project, we believe that the remaining network genes could be considered as potential candidates to develop new prognostic multigene expression profiles for breast cancer, to increase the sensitivity/specificity of current profiles, and/or as a means to potentially expand patient eligibility criteria.

Although conceptualized within the context of cognitive dysfunction, the identified genes would be ideal candidates for investigations of symptom variability in other disruptive and burdensome symptoms associated with breast cancer, its treatments, or both as well. The identified 21 gene set would be especially relevant for symptoms previously found to be associated with disease-related factor of breast cancer, such as fatigue, which has been predicted by tumor size and lymph involvement.[42,43] Moreover, the novel approach we employed to select candidate genes for investigations of variability in breast cancer symptoms can also be applied to other cancers that have biologically based commercially available prognostic multigene expression profiles, such as colon and prostate cancer, and associated symptoms.

Genetic variability within the proposed candidate gene set can be evaluated in a number of ways, including evaluation of polymorphisms, gene expression levels, protein levels, and epigenetic changes in both the host and tumor tissue. We recognize that symptom variability, especially at the time of diagnosis, may be driven by tumor gene expression, the consequences of tumor expression and related protein production on the rest of the body, and subsequent removal and treatment of the primary tumor and secondary sites. Thus, we recommend that future symptomatology studies focus on relationships between tumor gene expression/protein levels and symptom variability in cancer survivors. Significant results from tumor gene expression studies in particular, could greatly and directly expand the clinical utility of currently available prognostic multigene expression profiles. For instance, a modified version of the gene expression algorithms could potentially generate a range of Cognitive Decline Risk Scores or Fatigue Risk Scores,

based on the treatment regimen that is ultimately selected, other clinicopathologic tumor features, and baseline patient characteristics for each breast cancer survivor. However, relationships between host DNA and symptom variability are desirable as well because host DNA variation remains constant over time and is not tissue specific. Prediction of symptoms using host DNA would be especially advantageous for unusable tumor tissue, survivors with a clear prognosis and treatment regimen recommendation based on other clinicopathologic tumor features who would not be eligible for further prognostic multigene expression profile testing, or post hoc symptom prediction after tissue RNA has degraded and quantified gene expression level accuracy would be questionable.

However, we would like to acknowledge that this approach is not without limitations. Of particular interest, the multigene expression profiles from which candidate genes were selected, with the exception of the 70-gene breast cancer recurrence assay, all require positive breast cancer tumor estrogen or progesterone receptor status as an eligibility criterion. Consequently, genetic variation important for symptom variability in women with hormone receptor negative and triple-negative breast cancers may not be adequately captured in the prioritized gene set. In other words, the clinical applicability of breast cancer survivor symptom prediction may be limited to women with tumors that reflect the eligibility criteria of the five multigene expression profiles used to generate the candidate gene list. Additional candidate genes should be considered for other subsets of breast cancers not emphasized in the currently available multigene expression profiles.

Undoubtedly, the results of future investigations of symptom variability in breast cancer survivors based on disease characteristics at the cellular level, such as genetic variability in the high priority 21 gene set identified in this article, have the potential to substantially impact both the care of breast cancer survivors and the practice of health care providers alike, extending the clinical utility of prognostic multigene expression profiles for breast cancer and providing the patient and the provider with a means of weighing not only cancer prognosis and recurrence, but also the quality of life due to burdensome symptoms, into treatment decisions as well.

Acknowledgments

This project was supported by the Targeted Research and Academic Training for Nurses in Genomics (T32NR009759) training program and the Cognitive Function and Breast Cancer: Genomics and Disease Characteristics (F31NR014590) study. The authors acknowledge Dr Susan Cohen, PhD, FNP-BC, FAAN, and members of the University of Pittsburgh Cancer Survivorship Research Seminar for thoughtfully reviewing this manuscript. Portions of this project were presented at the International Society of Nurses in Genetics (ISONG) Silver Anniversary Conference in 2013.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Shi Q, Smith TG, Michonski JD, Stein KD, Kaw CK, Cleeland CS. Symptom burden in cancer survivors one year after diagnosis: a report form the American Cancer Society's Studies of Cancer Survivors. *Cancer*. 2012;117(12):2779–2790.
2. Burkett VS, Cleeland CS. Symptom burden in cancer survivorship. *J Cancer Surviv*. 2007;1(2):167–175.
3. Wu HS, Harden JK. Symptom burden and quality of life in survivorship: a review of the literature. *Cancer Nurs*. 2015;38(1):E29–E54.
4. Spiegel D, Giese-Davis J. Depression and cancer: mechanisms and disease progression. *Biol Psychiatr*. 2003;54(3):269–282.
5. Barsevick A, Frost M, Zwinderman A, Hall P, Halyard M; GENEQOL Consortium. I'm so tired: biological and genetic mechanisms of cancer-related fatigue. *Qual Life Res*. 2011;19(10):1419–1427.
6. Vardy J, Wefel JS, Ahles T, Tannock IF, Schagen SB. Cancer and cancer-therapy related cognitive dysfunction: an international perspective from the Venice cognitive workshop. *Ann Oncol*. 2008;19(4):623–629.
7. Falleti MG, Sanfilippo A, Maruff P, Weih L, Phillips KA. The nature and severity of cognitive impairment associated with adjuvant chemotherapy in women with breast cancer: a meta-analysis of the current literature. *Brain Cogn*. 2005;59(1):60–70.
8. Bender CM, Sereika SM, Brufsky AM, et al. Memory impairments with adjuvant anastrozole versus tamoxifen in women with early-stage breast cancer. *Menopause*. 2010;14(6):995–998.
9. Castellon SA, Ganz PA, Bower JE, Petersen L, Abraham L, Greendale GA. Neurocognitive performance in breast cancer survivors exposed to adjuvant chemotherapy and tamoxifen. *J Clin Exp Neuropsychol*. 2004;26(7):955–969.
10. Breckenridge LM, Bruns GL, Todd BL, Feuerstein M. Cognitive limitations associated with tamoxifen and aromatase inhibitors in employed breast cancer survivors. *Psychooncology*. 2012;53:43–53.
11. Lejbak L, Vrbancic M, Crossley M. Endocrine therapy is associated with low performance on some estrogen-sensitive cognitive tasks in postmenopausal women with breast cancer. *J Clin Exp Neuropsychol*. 2010;32(8):836–846.
12. Schilder CM, Seynaeve C, Beex LV, et al. Effects of tamoxifen and exemestane on cognitive functioning of postmenopausal patients with breast cancer: results from the neuropsychological side study of the tamoxifen and exemestane adjuvant multinational trial. *J Clin Oncol*. 2010;28(8):1294–1300.
13. Shilling V, Jenkins V, Fallowfield L, Howell T. The effects of hormone therapy on cognition in breast cancer. *J Steroid Biochem Mol Biol*. 2003;86(3–5):405–412.
14. Jenkins V, Shilling V, Fallowfield L, Howell A, Hutton S. Does hormone therapy for the treatment of breast cancer have a detrimental effect on memory and cognition?: a pilot study. *Psychooncology*. 2004;13(1):61–66.
15. Hermelink K, Untch M, Lux MP, et al. Cognitive function during neoadjuvant chemotherapy for breast cancer: results of a prospective, multicenter, longitudinal study. *Cancer*. 2007;109(9):1905–1913.

16. Wefel JS, Lenzi R, Theriault R, Buzdar AU, Cruickshank S, Meyers CA. "Chemobrain" in breast carcinoma?: a prologue. *Cancer*. 2004; 101(3):466–475.
17. Ahles TA, Avenue L, York N, Saykin AJ, Mcdonald BC, Kaufman PA. Cognitive function in breast cancer patients prior to adjuvant treatment. *Breast Cancer Res Treat*. 2011;110(1):143–152.
18. Bender CM, Sereika SM, Ryan CM, Brufsky AM, Puhalla S, Berga SL. Does lifetime exposure to hormones predict pretreatment cognitive function in women before adjuvant therapy for breast cancer? *Menopause*. 2013;20(9):1–8.
19. Schagen SB, Muller MJ, Boogerd W, Mellenbergh GJ, van Dam FS. Change in cognitive function after chemotherapy: a prospective longitudinal study in breast cancer patients. *J Natl Cancer Inst*. 2006;98(23):1742–1745.
20. Zhang Y, Schnabel CA, Schroeder BE, et al. Breast cancer index identifies early-stage estrogen receptor-positive breast cancer patients at risk for early- and late-distant recurrence. *Clin Cancer Res*. 2013;19(15):4196–4205.
21. Sgroi DC, Carney E, Zarrella E, et al. Prediction of late disease recurrence and extended adjuvant letrozole benefit by the HOXB13/IL17BR biomarker. *J Natl Cancer Inst*. 2013;105(14):1036–1042.
22. Sgroi DC, Sestak I, Cuzick J, et al. Prediction of late distant recurrence in patients with oestrogen-receptor positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population. *Lancet Oncol*. 2013;14(11):1067–1076.
23. Tutt A, Wang A, Rowland C, et al. Risk estimation of distant metastasis in node-negative, estrogen receptor-positive breast cancer patients using an RT-PCR based prognostic expression signature. *BMC Cancer*. 2008;8:339.
24. Genomic Health®, Inc. *Oncotype DX® Breast Cancer Assay Invasive Sales Aid*. Available from: http://breast-cancer.oncotypedx.com/en-US/Professional-Invasive/~/media/ODX-Breast/Images/Invasive/Invasive-SalesAid_11-2.pdf. Accessed November 3, 2016.
25. Genomic Health®, Inc. *Oncotype DX® Breast Cancer Assay DCIS Sales Aid*. Available from: http://breast-cancer.oncotypedx.com/en-US/Professional-DCIS/~/media/ODX-Breast/Files/Downloads/DCIS_Sales_Aid.pdf. Accessed November 13, 2015.
26. NanoString Technologies®, Inc. *Package Insert Prosigna® Breast Cancer Prognostic Gene Signature Assay*. Available from: http://prosigna.com/docs/Prosigna_Packet_Insert_US.pdf. Accessed October 5, 2015.
27. Agendia®. *Mammaprint® 70-Gene Breast Cancer Recurrence Assay*. Available from: http://www.agendia.com/healthcare-professionals/breast-cancer/mammaprint/. Accessed November 12, 2015.
28. Jerevall PL, Ma XJ, Li H, et al. Prognostic utility of HOXB13:IL17BR and molecular grade index in early-stage breast cancer patients from the Stockholm trial. *Br J Cancer*. 2011;104(11):1762–1769.
29. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004;351(27):2817–2826.
30. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *J Clin Oncol*. 2006;24(23):3726–3734.
31. Parker JS, Mullins M, Cheang MCU, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol*. 2009;27(8): 1160–1167.
32. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *J Natl Cancer Inst*. 2006;98(17):1183–1192.
33. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415(6871): 530–536.
34. Horn D, Zhou W, Trevisson E, et al. The conserved mitochondrial twin Cx9C protein Cmc2 Is a Cmc1 homologue essential for cytochrome coxidase biogenesis. *J Biol Chem*. 2010;285(20):15088–15099.
35. Vulsteke V, Beullens M, Boudrez A, et al. Inhibition of spliceosome assembly by the cell cycle-regulated protein kinase MELK and involvement of splicing factor NIPP1. *J Biol Chem*. 2004;279(10): 8642–8647.
36. Lin ML, Park JH, Nishidate T, Nakamura Y, Katagiri T. Involvement of maternal embryonic leucine zipper kinase (MELK) in mammary carcinogenesis through interaction with Bcl-G, a pro-apoptotic member of the Bcl-2 family. *Breast Cancer Res*. 2007;9(1):R17.
37. Beullens M, Vancauwenbergh S, Morrice N, et al. Substrate specificity and activity regulation of protein kinase MELK. *J Biol Chem*. 2005;280(48):40003–40011.
38. Lin YC, Chen CC, Cheng CJ, Yang RB. Domain and functional analysis of a novel breast tumor suppressor protein, SCUBE2. *J Biol Chem*. 2011;286(30):27039–27047.
39. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2015; 43(Database issue):D6-17.
40. Tian S, Roepman P, van't Veer LJ, Bernards R, de Snoo F, Glas AM. Biological functions of the genes in the mammaprint breast cancer profile reflect the hallmarks of cancer. *Biomark Insights*. 2010;5:129–138.
41. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646–674.
42. Goldstein D, Bennett BK, Webber K, et al. Cancer-related fatigue in women with breast cancer: outcomes of a 5-year prospective cohort study. *J Clin Oncol*. 2012;30(15):1805–1812.
43. Wielgus KK, Berger AM, Hertzog M. Predictors of fatigue 30 days after completing anthracycline plus taxane adjuvant chemotherapy for breast cancer. *Oncol Nurs Forum*. 2009;36(1):38–48.

APPENDIX I

LICENSE AGREEMENT FOR MANUSCRIPT #3

Dove Medical Press Ltd. LICENSE TERMS AND CONDITIONS

Mar 31, 2016

This is a License Agreement between University of Pittsburgh -- Theresa Koleck ("You") and Dove Medical Press Ltd. ("Dove Medical Press Ltd.") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Dove Medical Press Ltd., and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3839381284470
License date	Mar 31, 2016
Licensed content publisher	Dove Medical Press Ltd.
Licensed content title	Breast Cancer : Targets and Therapy
Licensed content date	Jan 1, 2009
Type of Use	Thesis/Dissertation
Requestor type	Academic institution
Format	Electronic
Portion	chapter/article
Title or numeric reference of the portion(s)	All
Title of the article or chapter the portion is from	Identification and prioritization of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level
Editor of portion(s)	N/A
Author of portion(s)	Koleck & Conley
Volume of serial or monograph.	8
Page range of the portion	None
Publication date of portion	2016
Rights for	Main product
Duration of use	Life of current and all future editions
Creation of copies for the disabled	no
With minor editing privileges	no
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional use	no
The lifetime unit quantity of	Up to 499

new product	
Made available in the following markets	universities
The requesting person/organization is:	Theresa Koleck
Order reference number	None
Author/Editor	Theresa Koleck
The standard identifier	KoleckETD
Title	Cognitive Function and Breast Cancer: Genomics and Disease Characteristics
Publisher	University of Pittsburgh
Expected publication date	Apr 2017
Estimated size (pages)	250
Total (may include CCC user fee)	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS
The following terms are individual to this publisher:

None

Other Terms and Conditions:
STANDARD TERMS AND CONDITIONS

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.
2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.
3. Scope of License; Limitations and Obligations.
3.1 All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.
3.2 General Payment Terms: You may pay by credit card or through an account with us

payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, Dept 001, P.O. Box 843006, Boston, MA 02284-3006. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3 Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4 In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5 Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6 User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and

expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. **Miscellaneous.**

8.1 User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

8.2 Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here:http://www.copyright.com/content/cc3/en/tools/footer/privacypolicy.html.

8.3 The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms

and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

8.4 No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

8.5 The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to info@copyright.com.

v 1.1

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

APPENDIX J

MANUSCRIPT #4: POLYMORPHISMS IN DNA REPAIR AND OXIDATIVE STRESS GENES ASSOCIATED WITH PRE-TREATMENT COGNITIVE FUNCTION IN BREAST CANCER SURVIVORS: AN EXPLORATORY STUDY

Koleck et al. SpringerPlus (2016) 5:422
DOI 10.1186/s40064-016-2061-4

SpringerPlus

RESEARCH Open Access

Polymorphisms in DNA repair and oxidative stress genes associated with pre-treatment cognitive function in breast cancer survivors: an exploratory study

Theresa A. Koleck[1], Catherine M. Bender[1], Susan M. Sereika[1,2,3], Adam M. Brufsky[4,5,6], Barry C. Lembersky[4,5], Priscilla F. McAuliffe[5,6,7], Shannon L. Puhalla[4,6], Priya Rastogi[4,6] and Yvette P. Conley[1,8*]

*Correspondence: yconley@pitt.edu
[1] School of Nursing, University of Pittsburgh, 3500 Victoria Street, Pittsburgh, PA 15261, USA
Full list of author information is available at the end of the article

Abstract

Purpose: The purpose of this exploratory candidate gene association study was to examine relationships between polymorphisms in oxidative stress and DNA repair genes and pre-adjuvant therapy cognitive function (CF) in postmenopausal women diagnosed with early stage-breast cancer.

Methods: Using a neuropsychological test battery, CF was assessed in 138 women diagnosed with breast cancer prior to initiation of adjuvant therapy and 81 age- and education-matched controls and summarized across eight composites. Participants were genotyped for 39 functional or tagging single nucleotide polymorphisms (SNPs) of select oxidative stress (*CAT*, *GPX1*, *SEPP1*, *SOD1*, and *SOD2*) and DNA repair (*ERCC2*, *ERCC3*, *ERCC5*, and *PARP1*) genes. Multiple linear regression was used to determine if the presence or absence of one or more minor alleles account for variability in CF composite scores. Based on regression findings from the analysis of individual SNPs, weighted multi-gene, multi-polymorphism genetic risk scores (GRSs) were calculated to evaluate the collective effect of possession of multiple protective and/or risk alleles.

Results: Each CF composite was significantly ($p < 0.05$) associated with one or more oxidative stress and DNA repair gene polymorphisms evaluated either by SNP main effects and/or SNP-by-prescribed breast cancer treatment group interactions. Each computed GRS was found to be significantly ($p < 0.001$) related to its corresponding CF composite. All associations were positive suggesting that as overall genetic protection increases, CF composite score increases (indicating better performance).

Conclusions: These findings suggest that genetic variation in the oxidative stress and DNA repair pathways may play an important role in pre-adjuvant therapy CF in breast cancer survivors.

Keywords: Breast neoplasms, Cognition, Genetics, Polymorphisms, Oxidative stress, DNA repair

Background

Pretreatment cognitive dysfunction has been well documented in women diagnosed with breast cancer (Wefel et al. 2008; Ahles et al. 2012); however, the mechanisms

underlying this phenomenon as well as the variability in the presence and severity of cognitive dysfunction experienced by breast cancer survivors remain largely unknown. One biologically plausible mechanism that may at least partially account for pretreatment cognitive dysfunction and the observed variability is variation in response to oxidative stress and DNA damage (Janelsins et al. 2012; Ahles and Saykin 2007; Vardy et al. 2008). Evidence continues to build that supports the role of increased oxidative stress, insufficient antioxidant mechanisms, and/or deficient response to DNA damage in brain aging and cognitive decline (Coppedè and Migliore 2010; Jeppesen et al. 2012; Lillenes et al. 2011). Furthermore, oxidative damage and diminished DNA repair capacity have been implicated in more extreme cognitive dysfunction phenotypes, including mild cognitive impairment and Alzheimer's disease (Bucholtz and Demuth 2013; Migliore et al. 2005; Jones et al. 1989).

The systemic environment and tumor microenvironment of a woman with breast cancer are characterized by increased, yet variable, levels of oxidative stress and DNA damage, with oxidative stress and subsequent DNA damage promoting breast cancer development and progression (Kang 2002; Jezierska-Drutel et al. 2013; Nourazarian et al. 2014). In one study of altered oxidative stress levels and breast cancer, Herrera et al. (2014) found evidence to support enhanced oxidative stress and reduced antioxidant defenses in plasma of postmenopausal women with primary ductal carcinomas of the breast at diagnosis compared to women 6 months post tumor removal and to healthy controls. Wang et al. also found increased levels of lipid peroxidation in breast cancer tissue but, in contrast to the previous study, upregulated antioxidant levels compared to tissue from healthy controls (Wang et al. 2014). In addition to being altered, research suggests that oxidative stress profiles are heterogeneous, differing between early and advanced stage breast cancers (Panis et al. 2012) and varying by tumor size and lymph node involvement (Saintot et al. 2002). In terms of DNA damage, chemotherapy naïve postmenopausal women with primary invasive ductal breast cancer were found to have higher basal levels of DNA damage and decreased DNA repair efficacy of peripheral blood lymphocytes (PBLs) compared to age-matched healthy women (Blasiak et al. 2004). Sanchez-Suarez et al. (2008) presented congruent findings: an assessment of PBLs from women with Stage 2 ductal carcinoma of the breast displayed higher DNA damage prior to initiation of adjuvant chemotherapy compared to age-matched healthy controls. Similar results were reported in a study of DNA damage and repair in PBLs in a heterogeneous sample of individuals (ages 1–59 years) with various cancer diagnoses compared to healthy controls (ages 22–50 years) with cells from cancer patients demonstrating higher levels of basal DNA damage. Considerable individual variation was also noted (Nadin et al. 2006).

The reported variability in oxidative stress and DNA damage profiles warrants investigation of genetic variation to account for differences in cognitive phenotypes of women diagnosed with breast cancer. Considering how increased oxidative stress and decreased DNA repair capacity impacts more extreme cognitive phenotypes as well as the vulnerability of the brain within the context of increased oxidative stress due to breast cancer, we hypothesize that variability in protection from oxidative damage and capacity to repair DNA may inform variability in the extent of cognitive dysfunction among breast cancer survivors. Thus, the purpose of this exploratory candidate gene association study

was to examine relationships between variation in genes involved in oxidative stress (*CAT*, *GPX1*, *SEPP1*, *SOD1*, and *SOD2*) and DNA repair (*ERCC2*, *ERCC3*, *ERCC5*, and *PARP1*) and pre-adjuvant therapy CF in postmenopausal women with early stage breast cancer. Furthermore, cumulative multi-gene, multi-polymorphism genetic risk scores (GRSs) were calculated to evaluate the collective effect of possessing multiple significant polymorphisms.

Methods

Study participants

Participants in this candidate gene association study were recruited from a larger parent study investigating the effect of the adjuvant aromatase inhibitor therapy, anastrozole, on changes in CF in postmenopausal women with breast cancer (Bender et al. 2015). The sample (N = 219) was comprised of 138 women diagnosed with breast cancer and 81 age- and education-matched controls with no personal history of breast cancer. All participants were no greater than 75 years of age, able to speak and read English, completed at least 8 years of education, and had no previous history of cancer, psychiatric illness, or neurologic disease/trauma at time of enrollment into the parent study. In addition, women with breast cancer had a diagnosis of estrogen receptor positive, early-stage breast cancer (stages 1, 2, or 3a) based on the Tumor, Nodes, Metastasis Classification of Malignant Tumors with no clinical evidence of distant metastases (Edge et al. 2010). This study was approved by the University of Pittsburgh Institutional Review Board. Informed consent was obtained from all individual participants included in both the parent and genetic ancillary study.

Evaluation of cognitive function

A battery of neuropsychological tests was used to assess cognitive function (CF). Women with cancer completed cognitive assessment after primary surgery but prior to initiation of adjuvant therapy. Control women completed the same cognitive assessment. The neuropsychological test battery was individually administered to study participants by trained research nurses. The selection of neuropsychological tests for the battery and reduction of individual neuropsychological test data into the eight following CF composites based on exploratory factor analysis have been described in detail previously (Bender et al. 2015):

1. Attention—Cambridge Neuropsychological Test Automated Battery (CANTAB) Rapid Visual Information Processing (Robbins et al. 1994)
2. Concentration—Digit Vigilance (Layfayette Clinical Insturments Company 1989)
3. Mental Flexibility—Delis Kaplan Executive Function System Color-Word Interference (Delis et al. 2001)
4. Executive Function—CANTAB Stockings of Cambridge (Robbins et al. 1994) and CANTAB Spatial Working Memory (Robbins et al. 1994)
5. Psychomotor Speed—Grooved Pegboard (Klove 1963) and Digit Symbol Substitution (Wechsler 1998)
6. Verbal Memory—Rey Auditory Verbal Learning (Rey 1964), Verbal Fluency Test, and Rivermead Story (Cockburn and Smith 1993)

7. Visual Memory—CANTAB Paired Associates Learning (Robbins et al. 1994) and Rey Complex Figure (Osterrieth 1944)
8. Visual Working Memory—CANTAB Stockings of Cambridge (Robbins et al. 1994) and Rey Complex Figure (Osterrieth 1944)

Covariate assessment

In order to control for the potential effects of age, intelligence, mood, and pain on CF, age, estimated verbal intelligence (National Adult Reading Test-Revised) (Nelson 1981), and levels of depressive symptoms (Beck Depression Inventory-II) (Beck et al. 1996), anxiety (POMS tension-anxiety subscale) (McNair et al. 1992), fatigue (POMS fatigue-inertia subscale) (McNair et al. 1992), and pain (Brief Pain Inventory) (Cleeland 1989) were also assessed. All participants in this study had complete covariate/confounder information.

SNP selection and genotyping

Functional polymorphisms for five candidate oxidative stress genes (Catalase, *CAT*; Glutathione Peroxidase 1, *GPX1*; Selenoprotein P, Plasma 1, *SEPP1*; Superoxide Dismutase 1, Soluble, *SOD1*; and Superoxide Dismutase 2, Mitochondrial, *SOD2*) and four candidate DNA repair genes (Excision Repair Cross-Complementation Group 2, *ERCC2*; Excision Repair Cross-Complementation Group 3, *ERCC3*; Excision Repair Cross-Complementation Group 5, *ERCC5*; and Poly (ADP-ribose) Polymerase 1, *PARP1*) were identified from the literature (Hamanishi et al. 2004; Valenti et al. 2004; Islam et al. 2007; Jiang et al. 2001; De Haan et al. 1998; Meplan et al. 2007; Spencer et al. 2008; Hooker et al. 2007; Mizutani 2007; Lockett et al. 2004). When a functional polymorphism was not identified or did not fully represent all of the variability in the gene, tagging SNPs were selected using the Phase III HapMap database. Criteria for selecting tagging SNPs included: R^2 of $\geq$0.8; minor allele frequency $\geq$20 %; and selected for Caucasian and African ancestry, which represents parent study subjects. In total, 39 functional or tagging SNPs were selected for evaluation (Table 1).

Genetic samples were collected from June 2008 to May 2014. Three milliliters of whole blood or two milliliters of saliva were obtained for genotyping. DNA was extracted from PBLs using a simple salting out procedure or from saliva utilizing the protocol and reagents supplied with the Oragene DNA collection kits (DNA Genotek Inc 2012; Miller et al. 1988). Genotypes were determined using either an iPLEX MassARRAY multiplex assay platform (Sequenom, San Diego, CA) or a TaqMan allele discrimination platform (Thermo Fisher Scientific Inc., Waltham, MA). Genotypes were double called by individuals blinded to subject phenotypes and discrepancies addressed by reviewing raw data or re-genotyping. Participant genotypes were classified for data analysis based on the presence (i.e., homozygous variant genotype plus heterozygous genotype) or absence (i.e., wildtype genotype) of the minor allele.

Statistical analysis

Analyses were performed using IBM® SPSS® Statistics Version 23 (IBM Corp., Armonk, NY). A detailed descriptive analysis of all data was first performed to identify any anomalies prior to modeling. Each SNP was tested for Hardy–Weinberg equilibrium

Table 1 Candidate DNA repair and oxidative stress genes and associated SNPs

DNA repair genes		Oxidative stress genes	
ERCC2	*ERCC5*	*CAT*	*SEPP1*
rs13181	rs11069498	rs1001179[a]	rs230819
rs1799786	rs2296147	rs10488736	rs28919892
rs1799787	rs2296148[a]	rs2179625	rs3877899[a]
rs238406	rs4150355	rs511895	
rs238416	rs4150360	rs525938	
rs3916874	rs4771436	rs566979	*SOD1*
rs50871	rs751402	rs769214[a]	rs1041740
rs50872	rs873601		
ERCC3	*PARP1*	*GPX1*	*SOD2*
rs2134794	rs1136410[a]	rs1050450[a]	rs4880[a]
rs4150402	rs2271347		rs5746136
rs4150407	rs3219058		rs8031
rs4150477	rs3219090		

SNP single nucleotide polymorphism

[a] Functional polymorphism

using a Chi square goodness-of-fit test. To account for the heterogeneity of breast cancer tumors, women diagnosed with breast cancer were further classified prior to analysis using prescribed treatment regimen as a surrogate for disease characteristics, such as disease stage and aggressiveness. Subsequently, the analysis featured two groups of women diagnosed with breast cancer, Group A (prescribed chemotherapy followed by anastrozole, n = 55) and Group B (prescribed anastrozole alone, n = 83), as well as the reference, healthy age- and education-matched control group (n = 81). Hierarchical multiple linear regression modeling was employed to estimate relationships between individual SNPs and each CF composite score. Both main SNP effect only and SNP-by-group interaction models were fitted. In all models, the prescribed treatment groups, Group A and Group B, were compared to the reference, control group. Likewise, possession of one or more minor alleles (i.e., homozygous variant genotype plus heterozygous genotype) was compared to the reference, wildtype genotype. All models were adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain at study entry. Regression diagnostics were examined for each model. Potentially influential cases were identified and sensitivity analyses were conducted to evaluate the robustness of findings. In order to retain cases found to be influential due to extreme CF scores, scores were modified to be less extreme but still the highest/lowest CF score(s) for the affected composite. Unstandardized regression b-coefficients were obtained and tested at a two-tailed significance level of 0.05.

GRSs were then calculated for each participant to evaluate the collective effect of multiple DNA repair and oxidative stress polymorphisms on CF composite scores. Separate GRSs were calculated for each CF composite. SNP minor alleles found to be significantly ($p < 0.05$) negatively or positively associated with CF composites in the individual main effect only and/or interaction effect models were utilized in GRS calculations. In order to assign greater risk/protection to alleles with stronger associations, a weighted method was employed. Unstandardized regression b-coefficients from the individual SNP models were multiplied by 0 (absence) or 1 (presence) based on a participant's genotype and

prescribed treatment group membership and then summed. For example, the equation to calculate the verbal memory GRS would be as follows:

$$\begin{aligned}\text{Verbal Memory GRS} &= (-.346 * \textit{CAT}\text{rs}566979 - \text{G}) \\ &+ (.282 * \textit{CAT}\text{rs}566979 - \text{G} * \text{GroupA}) + (.387 * \textit{CAT}\text{rs}566979 - \text{G} * \text{GroupB}) \\ &+ (-.129 * \textit{ERCC}5\text{rs}11069498 - \text{G}) + (.536 * \textit{ERCC}5\textit{rs}11069498 - \text{G} * \text{GroupA}) \\ &+ (.190 * \textit{ERCC}5\textit{rs}11069498 - \text{G} * \text{GroupB}) + (-.075 * \textit{ERCC}5\text{rs}751402 - \text{C}) \\ &+ (.486 * \textit{ERCC}5\textit{rs}751402 - \text{C} * \text{GroupA}) + (.255 * \textit{ERCC}5\text{rs}751402 - \text{C} * \text{GroupB}) \\ &+ (-.074 * \textit{ERCC}5\text{rs}4150360 - \text{T}) + (.568 * \textit{ERCC}5\textit{rs}4150360 - \text{T} * \text{GroupA}) \\ &+ (.104 * \textit{ERCC}5\text{rs}4150360 - \text{T} * \text{GroupB})\end{aligned}$$

Thus, a participant prescribed chemotherapy plus anastrozole (Group A) who possessed the minor alleles for *CAT*rs566979 and *ERCC5*rs4150360 would have a verbal memory GRS of 0.43 calculated as follows:

$$\begin{aligned}&\text{Verbal Memory GRS} \\ &= (-.346 * 1) + (.282 * 1 * 1) + (.387 * 1 * 0) + (-.129 * 0) + (.536 * 0 * 1) \\ &+ (.190 * 0 * 0) + (-.075 * 0) + (.486 * 0 * 1) + (.255 * 0 * 0) + (-.074 * 1) \\ &+ (.568 * 1 * 1) + (.104 * 1 * 0) = 0.43\end{aligned}$$

A lower GRS indicates greater genetic risk for poorer CF and a higher GRS indicates greater genetic protection. Please note that if influential observations were identified by the sensitivity analysis, b-coefficients from the models with modified extreme CF scores were used. The unique contributions of GRSs in explaining the variance in CF composite scores were evaluated in the final block in a hierarchical multiple linear regression model, adjusted for age, estimated verbal intelligence, and levels of depressive symptoms, anxiety, fatigue, and pain. Participants missing genetic data necessary for completion of a GRS calculation were not included in the GRS analysis.

Results

Genotyping rates of the 39 SNPs ranged from 85.5 to 100 %. When considering the entire cohort (cases and controls), six SNPs were not in Hardy–Weinberg equilibrium: *ERCC2*rs1799786 ($\chi^2 = 4.77$, $p = 0.029$), *ERCC2*rs238416 ($\chi^2 = 3.92$, $p = 0.048$), *ERCC2*rs50871 ($\chi^2 = 4.37$, $p = 0.037$), *PARP1*rs1136410 ($\chi^2 = 4.78$, $p = 0.029$), *PARP1*rs3219090 ($\chi^2 = 6.04$, $p = 0.014$), and *SEPP1*rs28919892 ($\chi^2 = 4.29$, $p = 0.038$). Of these six SNPs, only one, ERCC2rs238416 ($\chi^2 = 4.29$, $p = 0.038$) was not in Hardy–Weinberg equilibrium when considering the control group alone. This deviation is most likely due to lack of random sampling from the population for both the cases and controls. Group-wise comparisons of participant characteristics revealed that study groups differed statistically, but not clinically significantly by age and estimated verbal intelligence (Table 2).

Results from the individual SNP variant regression analyses are reported in the table found in Additional file 1. Individual polymorphisms significantly ($p < 0.05$) associated with a particular CF composite by a SNP main effect and/or SNP-by-group interaction effect are listed in Table 3. A selection of results from the individual SNP analysis has been highlighted by CF composite in the text to follow; please note that all reported b-coefficients indicate the magnitude and direction of possession of one or more minor

Table 2 Participant characteristics (N = 219)

Characteristic (measure)	Group A (n = 55) *Mean* ± *SD* or n (%)	Group B (n = 83) *Mean* ± *SD* or n (%)	Healthy controls (n = 81) *Mean* ± *SD* or n (%)	F or χ^2 test statistic p value
Age (years)	58.76 ± 5.47	62.47 ± 5.96	60.06 ± 6.08	<.001*
Education (years)	15.67 ± 2.78	14.95 ± 3.06	14.84 ± 2.91	.232
Estimated verbal intelligence (NART-R)	108.94 ± 8.87	107.04 ± 8.84	114.72 ± 7.84	<.001*
Depression (BDI-II)	5.24 ± 6.61	4.60 ± 4.65	4.83 ± 5.52	.760
Anxiety (POMS tension-anxiety subscale)	9.61 ± 6.14	6.97 ± 4.65	6.61 ± 5.63	.004*
Fatigue (POMS fatigue-inertia subscale)	5.11 ± 5.33	5.84 ± 6.35	5.74 ± 5.99	.759
Pain (BPI)	1.47 ± 1.96	1.55 ± 2.27	0.94 ± 2.07	.144
Marital status, married	38 (69.1)	54 (65.1)	46 (56.8)	.306
Number of children	1.75 ± 1.22	2.05 ± 1.39	2.12 ± 1.53	.283
Race, Caucasian	52 (94.5)	81 (97.6)	75 (92.6)	.337
Cancer stage				
Stage 1	25 (45.5)	69 (83.1)	–	–
Stage 2a	19 (34.5)	12 (14.5)	–	–
Stage 2b	5 (9.1)	2 (2.4)	–	–
Stage 3a	6 (10.9)	0 (0.0)	–	–

SD, standard deviation; Group A, prescribed chemotherapy plus anastrozole; Group B, prescribed anastrozole alone; NART-R, National Adult Reading Test-Revised; BDI-II, Beck Depression Inventory-II, POMS, Profile of Mood States; BPI, Brief Pain Inventory. One-way ANOVAs utilized to compare study cohort means of continuous variables. Pearson's Chi square tests of independence used to examine the general associations between categorical variables

* $p < .05$

alleles on CF. For attention, possession of one or more *ERCC3*rs2134794 (b = −0.309, p = 0.010) or *ERCC5*rs873601 (b = −0.288, p = −0.015) minor alleles was associated with poorer performance regardless of group membership. SNP main effects also influenced mental flexibility, psychomotor speed, and concentration performance. For mental flexibility, a number of significant SNP main effects were observed over multiple oxidative stress and DNA repair genes: *ERCC2*rs13181 (b = −0.179, p = 0.031), *ERCC3*rs4150407 (b = 0.234, p = 0.016), *ERCC3*rs4150477 (b = 0.190, p = 0.038), *PARP1*rs2271347 (b = 0.202, p = 0.034), *SEPP1*rs230819 (b = 0.255, p = 0.018), and *SOD1*rs1041740 (b = 0.254, p = 0.006). Significant SNP main effects were also observed for psychomotor speed: *CAT*rs511895 (b = 0.237, p = 0.031), *CAT*rs769214 (b = −0.421, p = 0.020), *ERCC5*rs11069498 (b = −0.236, p = 0.044), *ERCC5*rs751402 (b = −0.224, p = 0.050), *ERCC5*rs873601 (b = −0.227, p = 0.037), and *SEPP1*rs3877899 (b = −0.327, p = 0.005). For concentration, possession of one or more minor alleles for every *SOD2* polymorphism evaluated, *SOD2*rs4880 (b = −0.303, p = 0.024), *SOD2*rs5746136 (b = −0.257, 0.023), *SOD2*rs8031 (b = −0.332, p = 0.011), contributed to poorer concentration performance regardless of group membership. In addition, the combination of Group B membership and possession of one or more *ERCC2*rs3916874 (b = 0.533, p = 0.050), *ERCC2*rs50872 (b = −0.882, p = 0.001), *ERCC3*rs4150407 (b = 0.546, p = 0.047), or *ERCC5*rs2296147 (b = 0.585, p = 0.043) minor alleles contributed

Table 3 Genetic risk score (GRS) and cognitive function composite regression analysis results

Composite cognitive function composite	Gene-SNP used in GRS calculation	Minor allele	Wildtype reference allele	b_{GRS}	Model R^2	R^2 change for GRS
Attention[a] (n = 214)	*ERCC3*-rs2134794	C	A	1.003*	0.236	0.048
	ERCC5-rs873601	G	A			
Concentration (n = 206)	*ERCC2*-rs3916874	C	G	0.619*	0.218	0.150
	ERCC2-rs50872	T	C			
	ERCC3-rs2134794	C	A			
	ERCC3-rs4150407	G	A			
	ERCC5-rs2296147	C	T			
	SOD2-rs4880	T	C			
	SOD2-rs5746136	A	G			
	SOD2-rs8031	T	A			
Executive function[a] (n = 215)	*ERCC3*-rs2134794	C	A	0.535*	0.299	0.075
	ERCC3-rs4150407	G	A			
	ERCC3-rs4150477	T	C			
	ERCC5-rs2296147	C	T			
	PARP1-rs2271347	A	G			
	PARP1-rs3219058	A	G			
Mental flexibility[a] (n = 198)	*ERCC2*-rs13181	G	T	0.669*	0.342	0.094
	ERCC3-rs4150407	G	A			
	ERCC3-rs4150477	T	C			
	PARP1-rs2271347	C	T			
	SEPP1-rs230819	A	C			
	SEPP1-rs3877899	A	G			
	SOD1-rs1041740	T	C			
Psychomotor speed[a] (n = 186)	*CAT*-rs511895	A	G	0.741*	0.288	0.126
	CAT-rs769214	G	A			
	ERCC5-rs11069498	G	A			
	ERCC5-rs2296148	T	C			
	ERCC5-rs751402	T	C			
	ERCC5-rs873601	G	A			
	SEPP1-rs3877899	A	G			
	SOD1-rs1041740	T	C			
Verbal memory (n = 214)	*CAT*-rs566979	G	T	0.567*	0.289	0.049
	ERCC5-rs11069498	G	A			
	ERCC5-rs4150360	C	T			
	ERCC5-rs751402	T	C			
Visual memory[a] (n = 178)	*CAT*-rs1001179	A	G	0.691*	0.260	0.118
	CAT-rs525938	G	A			
	CAT-rs566979	G	T			
	CAT-rs769214	G	A			
	ERCC5-rs751402	T	C			
	GPX1-rs1050450	G	A			
Visual working memory[a] (n = 210)	*ERCC2*-rs1799787	T	C	0.766*	0.256	0.111
	ERCC5-rs11069498	G	A			
	ERCC5-rs4150355	T	C			
	ERCC5-rs4150360	C	T			
	ERCC5-rs873601	G	A			
	PARP1-rs2271347	C	T			

SNP single nucleotide polymorphism, *GRS* genetic risk score. All regression models are adjusted for age, estimated verbal intelligence, and levels of depression, anxiety, fatigue, and pain

* $p < .001$

[a] GRS calculation based on b-coefficients from regression models with modified influential point values

positively or negatively to concentration scores. The combination of DNA repair gene variation and Group A membership was found to be associated with executive function performance with multiple significant SNP-by-Group A interaction effects observed: *ERCC3*rs2134794 (b = 0.470, p = 0.023), *ERCC3*rs4150407 (b = −0.466, p = 0.035), *ERCC3*rs4150477 (b = −0.417, p = 0.046), *ERCC5*rs2296147 (b = 0.477, p = 0.034), and *PARP1*rs2271347 (b = −0.589, p = 0.006). In contrast, the combination of Group B membership and possession of one or more *ERCC5*rs2296148 (b = 1.075, p = 0.024) or *SOD1*rs1041740 (b = −0.619, p = 0.015) minor alleles was associated with psychomotor speed scores. The combination of group membership and genetic variation was found to be important for all three memory-related cognitive composites as well. Specifically, the combination of Group A membership and possession of one or more minor alleles for: *ERCC5*rs11069498 contributed positively to verbal memory (b = 0.536, p = 0.034) and visual working memory (b = 0.629, p = 0.027) scores; *ERCC5*rs4150360 contributed positively to verbal memory (b = 0.568, p = 0.031) and visual working memory (b = 0.673, p = 0.023); and *ERCC5*rs751402 contributed positively to verbal memory (b = 0.486, p = 0.038) and visual memory (b = 0.499, p = 0.023). Additionally, the combination of Group A membership and *CAT* variation was found to be associated with visual memory: *CAT*rs1001179 (b = −0.512, p = 0.032), and *CAT*rs769214 (b = 0.480, p = 0.024). Two *CAT* SNP main effects, *CAT*rs525938 (b = −0.282, p = 0.049) and *CAT*rs566979 (b = −0.282, p = 0.049) were also observed with visual memory.

Each computed GRS was found to be significantly ($p < 0.001$) related to its respective CF composite (Table 3). All associations were found to be positive such that as GRSs increase, CF composite performance scores increase as well (Fig. 1).

Discussion

To our knowledge, this study represents the first investigation of relationships between oxidative stress and DNA repair gene variation and pre-adjuvant therapy CF in postmenopausal women diagnosed with early-stage breast cancer. Overall, our results revealed that performance for every CF composite was significantly ($p < 0.05$) associated with one or more oxidative stress and DNA repair gene polymorphisms by either SNP main effects (i.e., observed cognitive changes in both breast cancer survivors and healthy control women are associated with a certain genetic polymorphism) and/or SNP-by-group interaction effects (i.e., observed cognitive changes were associated with a certain combination of genetic polymorphism and prescribed treatment group).

Out of all the genes included in our investigation, variation in *ERCC5* appeared to influence cognitive performance most globally, with significant relationships noted between one or more *ERCC5* SNPs and every CF composite with the exception of mental flexibility. The function of *ERCC5* has been most widely investigated in xeroderma pigmentosum and DNA excision repair following UV-induced damage. More generally, *ERCC5* participates in nucleotide excision repair, encoding an endonuclease that makes 3′ incisions (US National Library of Medicine, National Institutes of Health 2014). *ERCC5* also decreases cellular oxidative burden, functioning as a cofactor for a DNA glycosylase that removes oxidized pyrimidines from DNA (US National Library of Medicine, National Institutes of Health 2014). Although rare, mutations in the *ERCC5* gene have also been associated with the development of Cockayne syndrome in combination

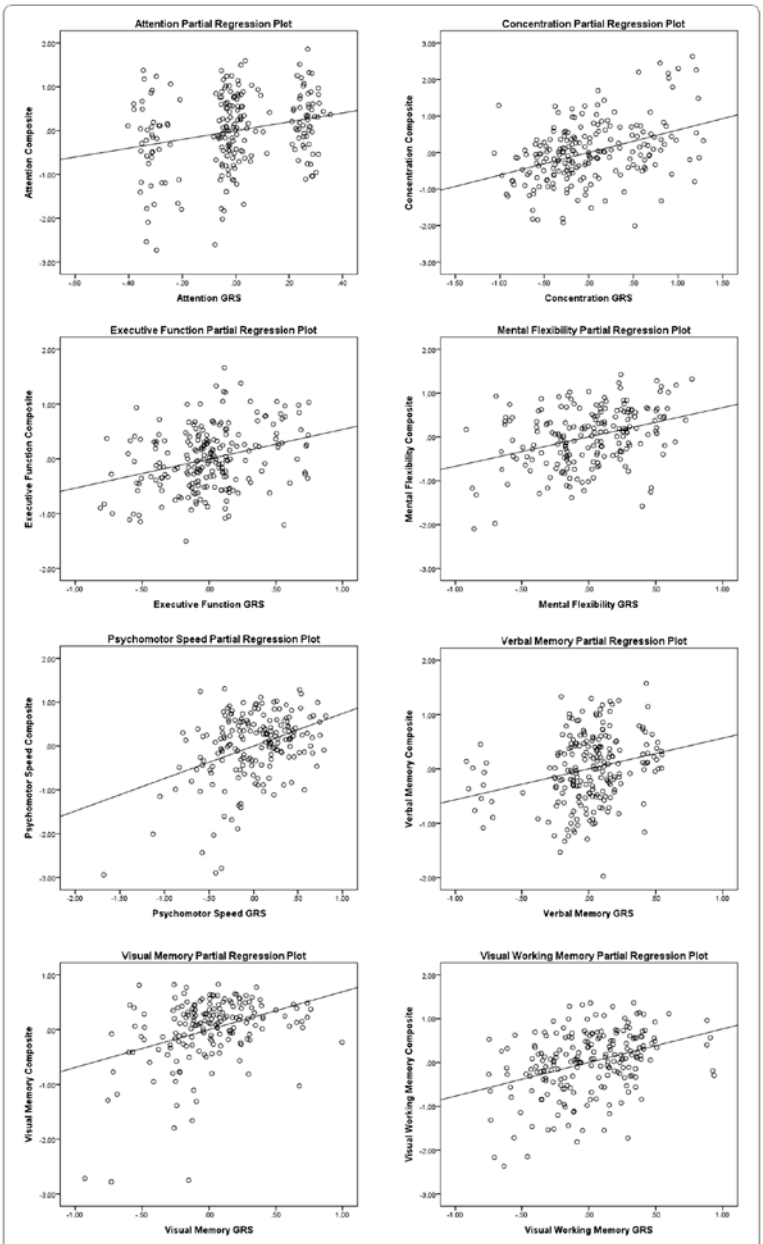


Fig. 1 Genetic risk score (GRS) partial regression plots. *Note* The partial regression plots were generated using IBM® SPSS® Statistics Version 23 (IBM Corp., Armonk, NY). *GRS* genetic risk score

with xeroderma pigmentosum (US National Library of Medicine, National Institutes of Health 2014). Characteristic features of Cockayne syndrome include impaired nervous system development and mental retardation, highlighting the critical role *ERCC5* plays in normal CF. Within the context of this study, we postulate that as cancer creates a cellular environment of increased oxidative stress and DNA damage, certain polymorphisms in *ERCC5* may decrease a survivor's ability to repair damage and remove reactive oxygen species, placing an already vulnerable brain at even higher risk for damage (Conroy et al. 2012; Joshi et al. 2005; Kasapović et al. 2010; Walker et al. 2012; Finkel and Holbrook 2000).

Another intriguing finding was the predominant negative effect of certain *SOD2* alleles (rs4880-T, rs5746136-A, and rs8031-T) on concentration performance within our study sample regardless of cancer diagnosis or prescribed treatment group. The *SOD2* gene encodes an enzyme, manganese-dependent superoxide dismutase, that confers cell protection by eliminating mitochondrial reactive oxygen species (NCBI Resource Coordinators 2015). The functional rs4880 alanine-to-valine (C > T) single amino acid change has been found to influence enzyme activity with valine (T) associated with reduced *SOD2* activity in human breast cancer and hepatoma cell lines (Sutton et al. 2005; McAteea and Yager 2010); contradictory findings have also been reported (Martin et al. 2009; Bastaki et al. 2006). Nevertheless, reduced *SOD2* expression has been implicated in a number of neurodegenerative disorders (Flynn and Melov 2013). Likewise, decreased *SOD2* mRNA and protein levels were found to be correlated with poorer memory, attention span, verbal fluency, and learning ability in a pooled sample of adults with recurrent depressive disorder and healthy controls (Talarowska et al. 2014). As our study was not designed to measure expression or protein levels of *SOD2*, we cannot expand upon how our results support or refute previous findings. While we found that possession of one or more *SOD2* rs4880-T alleles was associated with poorer CF in all study participants, the antioxidant properties of *SOD2* may have more impactful consequences for women with breast cancer throughout treatment with the introduction of adjuvant therapy regimens known to increase oxidative burden systemically. Thus, a remaining important question is if possession of one or more SOD2 rs4880-T alleles is also associated with greater risk for cognitive decline with therapy.

Our analyses also revealed a number of significant allele effects specific to the groups of women with breast cancer compared to control women without cancer. For example, significant SNP main effects and SNP-by-group interaction effects were observed for *PARP1* rs2271347 and executive function performance. While the SNP effect regression coefficient for possession of one or more minor alleles contributed positively to executive function ($b = 0.502$, $p < 0.001$), the SNP-by-group interaction regression coefficients for women in Group A (scheduled to received chemotherapy plus anastrozole) ($b = -0.589$, $p = 0.006$) and women in Group B (scheduled to receive anastrozole alone) ($b = -0.498$, $p = 0.006$) contributed negatively to the model, nullifying the main effect and contributing an overall negative input to executive function performance. Interactions, like the one presented, illustrate how not only genetic variation, but the combination of genetic variation and a breast cancer diagnosis can impact CF at diagnosis; alternatively, these findings highlight how factors that increase one's risk for development of cancer may also contribute to changes in CF.

In order to get a sense of the effect of the oxidative stress and DNA repair candidate pathway as a whole on pretreatment CF in women with breast cancer, we calculated weighted GRSs for each CF composite based on our individual SNP analysis. Similar to how a total score from an instrument (composed of individual items) summarizes a given concept, GRSs "summarize the potential multiple risk genetic influences into a single quantitative parameter and do not depend on single genetic variants" (Carreras-Torres et al. 2014). Instead of using a simple count method where each SNP contributes equally to risk calculations, we employed a weighted method in order to assign greater risk/protection to minor alleles with stronger associations (Lu et al. 2010). Fascinatingly, each computed GRS was significantly and positively associated to its respective CF composite. The amount of explained variance that each GRS contributed to its respective model was also notable, ranging from $R^2 = 0.048$ to 0.150. These findings not only point to the potential importance of oxidative stress response and DNA repair capacity to pretreatment changes in CF in breast cancer survivors, but, more broadly, to the value of evaluating the effect of multiple SNPs at the same time in association studies in general.

To better interpret our findings, a gene–gene pathway analysis, using Ingenuity® Pathway Analysis software (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity), of the nine candidate oxidative stress and DNA repair genes analyzed within this study was conducted. This analysis reiterated the interconnectedness of the genes and, consequently, endorsed evaluation of the collective effect of multiple SNPs from a single pathway simultaneously. Two unique networks were identified through our analysis (Additional file 2). The first network included *CAT*, *GPX1*, *PARP1*, *SOD1*, and *SOD2*. The main associated diseases and functions of this network were, not surprisingly, free radical scavenging, small molecule biochemistry, and neurological disease. The second network included *ERCC2*, *ERCC3*, *ERCC5*, and *SEPP1* and was associated with DNA replication, recombination, and repair, energy production, and nucleic acid metabolism. The pathway analysis also reminds that our study featured a limited number of candidate genes and that there are many additional oxidative stress and DNA repair genes that warrant further investigation.

While this study had a number of strengths, including mechanistic pathway-driven candidate gene selection, inclusion of a matched control group of women without a breast cancer diagnosis, assessment of SNP-by-prescribed treatment group interaction effects, and evaluation of the collective effect of multiple SNPs using weighted GRSs, limitations should also be acknowledged. To begin, the small study sample size, while appreciable, did not allow for detection of small effect sizes (i.e., $R^2 < .01$) often characteristic of genetic association studies. In addition, the small sample size did not allow for the evaluation of allelic dose–response relationships. Because the sample was comprised of postmenopausal women with hormone receptor positive, early-stage breast cancer who were primarily white and married, the generalizability to other more diverse populations and breast cancers is unknown. Findings from this study should be replicated in a larger, more diverse sample. In addition, differences in anesthesia exposure (or lack thereof for control women who did not undergo surgery) and its potential confounding cognitive effects were not controlled for in our analysis. Future studies and analyses should also focus on the collective effect of multiple oxidative stress and DNA repair gene variants on CF throughout and following completion of adjuvant therapy in women with breast cancer.

In conclusion, our goal in this study was to provide data on a possible biologic mechanism to account for variability in cognitive changes in breast cancer survivors. Results from this preliminary study reveal that genetic variation in the oxidative stress and DNA repair pathways appears to play an important role in CF in women with breast cancer prior to initiation of adjuvant therapy and give reason to investigate whether polymorphisms influence cognitive decline with therapy as well. In the future, evaluation of a panel of oxidative stress and DNA repair gene polymorphisms could offer healthcare providers a means of predicting which women diagnosed breast cancer are most at risk for poorer CF and candidates for additional interventions, such as antioxidant therapy.

Additional files

Additional file 1. Individual SNP and cognitive function regression analyses.
Additional file 2. Oxidative stress and DNA repair gene-gene networks generated by pathway analysis.

Abbreviations

BDI-II: Beck Depression Inventory-II; BPI: Brief Pain Inventory; CAT: catalase; CF: cognitive function; ERCC2: Excision Repair Cross-Complementation Group 2; ERCC3: Excision Repair Cross-Complementation Group 3; ERCC5: Excision Repair Cross-Complementation Group 5; GPX1: Glutathione Peroxidase 1; GRS: genetic risk score; NART-R: National Adult Reading Test-Revised; PARP1: Poly (ADP-ribose) Polymerase 1; PBL: peripheral blood leukocytes; POMS: Profile of Mood States; SD: standard deviation; SEPP1: Selenoprotein P, Plasma 1; SNP: single nucleotide polymorphism; SOD1: Superoxide Dismutase 1, Soluble; SOD2: Superoxide Dismutase 2, Mitochondrial.

Authors' contributions

Study conception and design: YPC, CMB; acquisition of data and samples: YPC, TAK, CMB, AMB, BCL, PFM, SLP, PR; data analyses: YPC, TAK, CMB, SMS; all authors participated in manuscript development. All authors read and approved the final manuscript.

Author details

[1] School of Nursing, University of Pittsburgh, 3500 Victoria Street, Pittsburgh, PA 15261, USA. [2] Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, 130 De Soto Street, Pittsburgh, PA 15261, USA. [3] Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, 130 De Soto Street, Pittsburgh, PA 15261, USA. [4] Division of Hematology/Oncology, Magee-Womens Hospital of University of Pittsburgh Medical Center (UPMC), 300 Halket Street, Pittsburgh, PA 15213, USA. [5] University of Pittsburgh Cancer Institute, 5150 Centre Avenue, Pittsburgh, PA 15232, USA. [6] School of Medicine, University of Pittsburgh, 3550 Terrace Street, Pittsburgh, PA 15261, USA. [7] Division of Breast Surgical Oncology, Magee-Womens Hospital of University of Pittsburgh Medical Center (UPMC), 300 Halket Street, Pittsburgh, PA 15213, USA. [8] Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, 130 De Soto Street, Pittsburgh, PA 15261, USA.

Acknowledgements

The authors would like to acknowledge and thank the women who participated in this study. This research was supported, in part, by the Oncology Nursing Society Foundation, National Institute of Nursing Research (F31NR014590, T32NR009759), National Cancer Institute (R01CA107408), and American Cancer Society (DSCN-14-076-01-SCN).

Competing interests

The authors declare that they have no competing interests.

Human and animal rights and informed consent

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the University of Pittsburgh Institutional Review Board. Informed consent was obtained from all individual participants included in this study. This article does not contain any studies with animals performed by any of the authors

Received: 17 March 2016 Accepted: 27 March 2016
Published online: 09 April 2016

References

Ahles TA, Saykin AJ (2007) Candidate mechanisms for chemotherapy-induced cognitive changes. Nat Rev Cancer 7:192–201. doi:10.1038/nrc2073

Ahles TA, Root JC, Ryan EL (2012) Cancer- and cancer treatment-associated cognitive change: an update on the state of the science. J Clin Oncol 30:3675–3686. doi:10.1200/JCO.2012.43.0116

Bastaki M, Huen K, Manzanillo P, Chande N, Chen C, Balmes JR, Tager IB, Holland N (2006) Genotype–activity relationship for Mn-superoxide dismutase, glutathione peroxidase 1 and catalase in humans. Pharmacogenet Genomics 16:279–286. doi:10.1097/01.fpc.0000199498.08725.9c

Beck AT, Steer RA, Brown GK (1996) Beck Depression Inventory-II. The Psychological Corporation, San Antonio

Bender CM, Merriman JD, Gentry AL, Ahrendt GM, Berga SL, Brufsky AM, Casillo FE, Dailey MM, Erickson KI, Kratofil FM, McAuliffe PF, Rosenzweig MQ, Ryan CM, Sereika SM (2015) Patterns of change in cognitive function with anastrozole therapy. Cancer 121:2627–2636. doi:10.1002/cncr.29393

Blasiak J, Arabski M, Krupa R, Wozniak K, Rykala J, Kolacinska A, Morawiec Z, Drzewoski J, Zadrozny M (2004) Basal, oxidative and alkylative DNA damage, DNA repair efficacy and mutagen sensitivity in breast cancer. Mutat Res 554:139–148. doi:10.1016/j.mrfmmm.2004.04.001

Bucholtz N, Demuth I (2013) DNA-repair in mild cognitive impairment and Alzheimer's disease. DNA Repair (Amst) 12:811–816. doi:10.1016/j.dnarep.2013.07.005

Carreras-Torres R, Kundu S, Zanetti D, Esteban E, Via M, Moral P (2014) Genetic Risk score of NOS gene variants associated with myocardial infarction correlates with coronary incidence across Europe. PLoS ONE 9:e96504. doi:10.1371/journal.pone.0096504

Cleeland CS (1989) Measurement of pain by subjective report. In: Chapman CR, Loeser JD (eds) Advances in pain research and therapy, vol 12. Raven Press, New York, pp 391–403

Cockburn J, Smith PT (1993) Correlates of everyday memory among residents of Part III homes. Br J Clin Psychol 32:75–77

Conroy SK, McDonald BC, Smith DJ, Moser LR, West JD, Kamendulis LM, Klaunig JE, Champion VL, Unverzagt FW, Saykin AJ (2012) Alterations in brain structure and function in breast cancer survivors: effect of post-chemotherapy interval and relation to oxidative DNA damage. Breast Cancer Res Treat. doi:10.1007/s10549-012-2385-x

Coppedè F, Migliore L (2010) DNA repair in premature aging disorders and neurodegeneration. Curr Aging Sci 3:3–19. doi:10.2174/1874609811003010003#sthash.QosRTCeY.dpuf

De Haan JB, Griffiths P, Kelner M, Shea RDO, Cheung NS, Bronson RT, Silvestro MJ, Wild S, Zheng SS, Beart PM, Hertzog PJ, Kola I (1998) Mice with a homozygous null mutation for the most abundant glutathione peroxidase, Gpx1, show increased Susceptibility to the oxidative stress-inducing agents paraquat and hydrogen peroxide. J Biol Chem 273:22528–22536. doi:10.1074/jbc.273.35.22528

Delis DC, Kaplan E, Kramer JH (2001) Delis-Kaplan (D-KEFS) executive function system, examiners manual. The Psychological Corporation, San Antonio

DNA Genotek Inc (2012) Laboratory protocol for manual purification of DNA from whole sample. DNA Genotek Inc, Ottawa, pp 1–8

Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A (2010) AJCC Cancer Staging Manual, 7th edn. Springer, New York, NY

Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of ageing. Nature 408:239–247. doi:10.1038/35041687

Flynn JM, Melov S (2013) SOD2 in mitochondrial dysfunction and neurodegeneration. Free Radic Biol Med 62:4–12. doi:10.1016/j.freeradbiomed.2013.05.027

Hamanishi T, Furuta H, Kato H, Doi A, Tamai M, Shimomura H, Sakagashira S, Nishi M, Sasaki H, Sanke T, Nanjo K (2004) Functional variants in the glutathione peroxidase-1 (GPx-1) gene are associated with increased intima-media thickness of carotid arteries and risk of macrovascular diseases in Japanese type 2 diabetic patients. Diabetes 53:2455–2460. doi:10.2337/diabetes.53.9.2455

Herrera ACS, Victorino VJ, Campos FC, Verenitach BD, Lemos LT, Aranome AMF, Oliveira SR, Cecchini AL, Simão ANC, Abdelhay E, Panis C, Cecchini R (2014) Impact of tumor removal on the systemic oxidative profile of patients with breast cancer discloses lipid peroxidation at diagnosis as a putative marker of disease recurrence. Clin Breast Cancer 14:451–459. doi:10.1016/j.clbc.2014.05.002

Hooker S, Bonilla C, Akereyeni F, Ahaghotu C, Kittles R (2007) NAT2 and NER genetic variants and sporadic prostate cancer susceptibility in African Americans. Prostate Cancer Prostatic Dis. 11:349–356. doi:10.1038/sj.pcan.4501027

Islam T, Mcconnell R, Gauderman WJ, Avol E, Peters JM, Gilliland FD (2007) Ozone, oxidant defense genes, and risk of asthma during adolescence. Am J Respir Crit Care Med 177:388–395. doi:10.1164/rccm.200706-863OC

Janelsins MC, Kohli S, Mohile SG, Usuki K, Ahles T, Morrow GR (2012) An update on cancer- and chemotherapy-related cognitive dysfunction. Semin Oncol 38:431–438. doi:10.1053/j.seminoncol.2011.03.014.AN

Jeppesen D, Bohr V, Stevnsner T (2012) DNA repair deficiency in neurodegeration. Prog Neurobiol 94:166–200. doi:10.1016/j.pneurobio.2011.04.013.DNA

Jezierska-Drutel A, Rosenzweig S, Neumann C (2013) Role of oxidative stress and the microenvironment in breast cancer development and progression. Adv Cancer Res 119:107–125. doi:10.1016/B978-0-12-407190-2.00003-4.Role

Jiang Z, Akey JM, Shi J, Xiong M, Wang Y, Shen Y, Xu X, Chen H, Wu H, Xiao J, Lu D, Huang W, Jin L (2001) A polymorphism in the promoter region of catalase is associated with blood pressure levels. Hum Genet 109:95–98. doi:10.1007/s004390100553

Jones S, Nee L, Sweet L, Polinsky R, Bartlett J, Bradley W, Robinson S (1989) Decreased DNA repair in familial Alzheimer's disease. Mutat Res 219:247–255. doi:10.1016/0921-8734(89)90007-6

Joshi G, Sultana R, Tangpong J, Cole MP, St Clair DK, Vore M, Estus S, Butterfield DA (2005) Free radical mediated oxidative stress and toxic side effects in brain induced by the anti cancer drug adriamycin: insight into chemobrain. Free Radic Res 39:1147–1154. doi:10.1080/10715760500143478

Kang D (2002) Oxidative stress, DNA damage, and breast cancer. AACN Clin Issues 13:540–549. doi:10.1097/00044067-200211000-00007

Kasapović J, Pejić S, Stojiljković V, Todorović A, Radošević-Jelić L, Saičić ZS, Pajović SB (2010) Antioxidant status and lipid peroxidation in the blood of breast cancer patients of different ages after chemotherapy with 5-fluorouracil, doxorubicin and cyclophosphamide. Clin Biochem 43:1287–1293. doi:10.1016/j.clinbiochem.2010.08.009

Klove H (1963) Clinical neuropsychology. The Medical Clinics, North America

Layfayette Clinical Insturments Company (1989) Layfayette Clinical Repeatable Neuropsychological Battery
Lillenes MS, Espeseth T, Støen M, Lundervold AJ, Frye SA, Rootwelt H, Reinvang I, Tønjum T (2011) DNA base excision repair gene polymorphisms modulate human cognitive performance and decline during normal life span. Mech Ageing Dev 132:449–458. doi:10.1016/j.mad.2011.08.002
Lockett KL, Hall MC, Xu J, Zheng SL, Berwick M, Chuang S, Clark PE, Cramer SD, Lohman K, Hu JJ (2004) The ADPRT V762A genetic variant contributes to prostate cancer susceptibility and deficient enzyme function. Cancer Res 64:6344–6348. doi:10.1158/0008-5472.CAN-04-0338
Lu Y, Feskens EJM, Boer JMA, Imholz S, Verschuren WMM, Wijmenga C, Vaarhorst A, Slagboom E, Müller M, Dollé MET (2010) Exploring genetic determinants of plasma total cholesterol levels and their predictive value in a longitudinal study. Atherosclerosis 213:200–205. doi:10.1016/j.atherosclerosis.2010.08.053
Martin RCG, Li Y, Liu Q, Jensen NS, Barker DF, Doll MA, Hein DW (2009) Manganese superoxide dismutase V16A single-nucleotide polymorphism in the mitochondrial targeting sequence is associated with reduced enzymatic activity in cryopreserved human hepatocytes. DNA Cell Biol 28:3–7. doi:10.1089/dna.2008.0788
McAteea BL, Yager JD (2010) Manganese superoxide dismutase: effect of the ala16 val polymorphism on protein, activity, and mRNA levels in human breast cancer cell lines and stably transfected mouse embryonic fibroblasts. Mol Cell Biochem 335:107–118. doi:10.1007/s11010-009-0247-6
McNair D, Lorr M, Droppleman LF (1992) EdITS manual for the profile of mood states. EdITS/Educational and Industrial Testing Service, San Diego
Meplan C, Crosley LK, Nicol F, Beckett GJ, Howie AF, Hill KE, Horgan G, Mathers JC, Arthur JR, Hesketh JE (2007) Genetic polymorphisms in the human selenoprotein P gene determine the response of selenoprotein markers to selenium supplementation in a gender-specific manner (the SELGEN study). FASEB J 21:3063–3074. doi:10.1096/fj.07-8166com
Migliore L, Fontana I, Trippi F, Colognato R, Coppedè F, Tognoni G, Nucciarone B, Siciliano G (2005) Oxidative DNA damage in peripheral leukocytes of mild cognitive impairment and AD patients. Neurobiol Aging 26:567–573. doi:10.1016/j.neurobiolaging.2004.07.016
Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215. doi:10.1093/nar/16.3.1215
Mizutani H (2007) Mechanism of DNA damage and apoptosis induced by anticancer drugs through generation of reactive oxygen species. Yakugaku Zasshi 127:1837–1842
Nadin SB, Vargas-Roig LM, Drago G, Ibarra J, Ciocca DR (2006) DNA damage and repair in peripheral blood lymphocytes from healthy individuals and cancer patients: a pilot study on the implications in the clinical response to chemotherapy. Cancer Lett 239:84–97. doi:10.1016/j.canlet.2005.07.025
NCBI Resource Coordinators (2015) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 43:D6–D17. doi:10.1093/nar/gku1130
Nelson H (1981) Nelson Adult Reading Test (NART) manual. NFER-Nelson, Windsor
Nourazarian A, Kangari P, Salmaninejad A (2014) Roles of oxidative stress in the development and progression of breast cancer. Asian Pac J Cancer Prev 15:4745–4751. doi:10.7314/APJCP.2014.15.12.4745
Osterrieth PA (1944) Test of copying a complex figure: contribution to the study of perception and memory. Arch Psychol (Geneve) 30:206–356
Panis C, Victorino VJ, Herrera ACS, Freitas LF, De Rossi T, Campos FC, Simão AN, Barbosa DS, Pinge-Filho P, Cecchini R, Cecchini AL (2012) Differential oxidative status and immune characterization of the early and advanced stages of human breast cancer. Breast Cancer Res Treat 133:881–888. doi:10.1007/s10549-011-1851-1
Rey A (1964) L'examen psychologique dans les cas d'encephalopathie traumatique. Arch Psychol 1252:340–382
Robbins T, James M, Owen A, Sahakian B, McInnes L, Rabbitt P (1994) Cambridge neuropsychological test automated battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. Dementia 5:266–281. doi:10.1159/000106735
Saintot M, Mathieu-Daude H, Astre C, Grenier J, Simony-Lafontaine J, Gerber M (2002) Oxidant–antioxidant status in relation to survival among breast cancer patients. Int J Cancer. doi:10.1002/ijc.10099
Sánchez-Suárez P, Ostrosky-Wegman P, Gallegos-Hernández F, Peñarroja-Flores R, Toledo-García J, Bravo JL, Del Castillo ER, Benítez-Bribiesca L (2008) DNA damage in peripheral blood lymphocytes in patients during combined chemotherapy for breast cancer. Mutat Res 640:8–15. doi:10.1016/j.mrfmmm.2007.11.008
Spencer DMS, Bilardi RA, Koch TH, Post GC, Nafie JW, Kimura KI, Cutts SM, Phillips DR (2008) DNA repair in response to anthracycline-DNA adducts: a role for both homologous recombination and nucleotide excision repair. Mutat Res 638:110–121. doi:10.1016/j.mrfmmm.2007.09.005
Sutton A, Imbert A, Igoudjil A, Descatoire V, Cazanave S, Pessayre D, Degoul F (2005) The manganese superoxide dismutase Ala16Val dimorphism modulates both mitochondrial import and mRNA stability. Pharmacogenet Genomics 15:311–319. doi:10.1097/01213011-200505000-00006
Talarowska M, Orzechowska A, Szemraj J, Su KP, Maes M, Galecki P (2014) Manganese superoxide dismutase gene expression and cognitive functions in recurrent depressive disorder. Neuropsychobiology 70:23–28. doi:10.1159/000363340
US National Library of Medicine, National Institutes of Health (2014) Genetics Home Reference: ERCC5. http://ghr.nlm.nih.gov/gene/ERCC5. Accessed 26 Jan 2016
Valenti L, Conte D, Piperno A, Dongiovanni P, Fracanzani AL, Fraquelli M, Vergani A, Gianni C, Carmagnola L, Fargion S (2004) The mitochondrial superoxide dismutase A16 V polymorphism in the cardiomyopathy associated with hereditary haemochromatosis. J Med Genet 41:946–950. doi:10.1136/jmg.2004.019588
Vardy J, Wefel JS, Ahles T, Tannock IF, Schagen SB (2008) Cancer and cancer-therapy related cognitive dysfunction: an international perspective from the Venice cognitive workshop. Ann Oncol 19:623–629. doi:10.1093/annonc/mdm500
Walker CH, Drew BA, Antoon JW, Kalueff AV, Beckman BS (2012) Neurocognitive effects of chemotherapy and endocrine therapies in the treatment of breast cancer: recent perspectives. Cancer Invest 30:135–148. doi:10.3109/07357907.2011.636116

Wang C, Yu J, Wang H, Zhang J, Wu N (2014) Lipid peroxidation and altered antioxidant status in breast adenocarcinoma patients. Drug Res (Stuttg) 64:690–692. doi:10.1055/s00341372580
Wechsler D (1998) The Wechsler memory scale-revised. Psychological Corporation, San Antonio
Wefel JS, Witgert ME, Meyers CA (2008) Neuropsychological sequelae of non-central nervous system cancer and cancer therapy. Neuropsychology 18:121–131. doi:10.1007/s11065-008-9058-x

BIBLIOGRAPHY

Abdel-Fatah, T. M., Powe, D. G., Ball, G., Lopez-Garcia, M. A., Habashy, H. O., Green, A. R., … Ellis, I. O. (2010). Proposal for a modified grading system based on mitotic index and Bcl2 provides objective determination of clinical outcome for patients with breast cancer. *Journal of Pathology*, *222*(4), 388–399. doi:10.1002/path.2775

Agoulnik, I. U., Tong, X. W., Fischer, D. C., Körner, K., Atkinson, N. E., Edwards, D. P., … Kieback, D. G. (2004). A germline variation in the progesterone receptor gene increases transcriptional activity and may modify ovarian cancer risk. *Journal of Clinical Endocrinology and Metabolism*, *89*(12), 6340–6347. doi:10.1210/jc.2004-0114

Ahles, T. A., & Saykin, A. J. (2007). Candidate mechanisms for chemotherapy-induced cognitive changes. *Nature Reviews Cancer*, *7*(3), 192–201. doi:10.1038/nrc2073

Ahles, T. A., Saykin, A. J., Furstenberg, C. T., Cole, B., Mott, L. A., Skalla, K., … Silberfarb, P. M. (2002). Neuropsychologic impact of standard-dose systemic chemotherapy in long-term survivors of breast cancer and lymphoma. *Journal of Clinical Oncology*, *20*(2), 485–493. doi:10.1200/JCO.20.2.485

Ahles, T. A., Avenue, L., York, N., Saykin, A. J., Mcdonald, B. C., & Kaufman, P. A. (2011). Cognitive function in breast cancer patients prior to adjuvant treatment. *Breast Cancer Research and Treatment*, *110*(1), 143–152. doi:10.1007/s10549-007-9686-5.Cognitive

Ahles, T. A., Root, J. C., & Ryan, E. L. (2012). Cancer- and cancer treatment – associated cognitive change: an update on the state of the science. *Journal of Clinical Oncology*, *30*(30), 3675–3686. doi:10.1200/JCO.2012.43.0116

Ahles, T. A., Saykin, A. J., Noll, W. W., Furstenberg, C. T., Guerin, S., Cole, B., & Mott, L. A. (2003). The relationship of APOE genotype to neuropsychological performance in long-term cancer survivors treated with standard dose chemotherapy. *Psycho-Oncology*, *12*(6), 612–619. doi:10.1002/pon.742

Ahles, T., Li, Y., McDonald, B., Schwartz, G., Kaufman, P., Tsongalis, G., … Saykin, A. (2011). Longitudinal assessment of cognitive changes associated with adjuvant treatment for breast cancer: the impact of APOE and smoking. *Psycho-Oncology*, *72*(2), 181–204. doi:10.1038/nature13314.A

Aleskandarany, M. A., Soria, D., Green, A. R., Nolan, C., Diez-Rodriguez, M., Ellis, I. O., &

Rakha, E. A. (2015). Markers of progression in early-stage invasive breast cancer: a predictive immunohistochemical panel algorithm for distant recurrence risk stratification. *Breast Cancer Research and Treatment*, *151*(2), 325–333. doi:10.1007/s10549-015-3406-3

American Cancer Society. (2014). Breast cancer: what are the key statistics about breast cancer? Retrieved from http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-key-statistics

Arriagada, R., Lê, M. G., Contesso, G., Guinebretière, J. M., Rochard, F., & Spielmann, M. (2002). Predictive factors for local recurrence in 2006 patients with surgically resected small breast cancer. *Annals of Oncology*, *13*(9), 1404–1413. doi:10.1093/annonc/mdf227

Bao, J., Yu, K. D., Jiang, Y. Z., Shao, Z. M., & Di, G. H. (2014). The effect of laterality and primary tumor site on cancer-specific mortality in breast cancer: a SEER population-based study. *PLoS ONE*, *9*(4). doi:10.1371/journal.pone.0094815

Batchelder, A. J., Gordon-Weeks, A. N., & Walker, R. A. (2009). Altered expression of anti-apoptotic proteins in non-involved tissue from cancer-containing breasts. *Breast Cancer Research and Treatment*, *114*(1), 63–69. doi:10.1007/s10549-008-9988-2

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-II*. San Antonio: The Psychological Corporation.

Bender, C. M., Merriman, J. D., Gentry, A. L., Ahrendt, G. M., Berga, S. L., Brufsky, A. M., … Sereika, S. M. (2015). Patterns of change in cognitive function with anastrozole therapy. *Cancer*, *121*(15), 2627–2636. doi:10.1002/cncr.29393

Bender, C. M., Sereika, S. M., Berga, S. L., Vogel, V. G., Brufsky, A. M., Paraska, K. K., & Ryan, C. M. (2006). Cognitive impairment associated with adjuvant therapy in breast cancer. *Psycho-Oncology*, *15*(5), 422–430. doi:10.1002/pon.964

Bender, C. M., Sereika, S. M., Brufsky, A. M., Ryan, C. M., Vogel, V. G., Rastogi, P., … Berga, S. L. (2010). Memory impairments with adjuvant anastrozole versus tamoxifen in women with early-stage breast cancer. *Menopause*, *14*(6), 995–998. doi:10.1097/gme.0b013e318148b28b.Memory

Bender, C. M., Sereika, S. M., Ryan, C. M., Brufsky, A. M., Puhalla, S., & Berga, S. L. (2013). Does lifetime exposure to hormones predict pretreatment cognitive function in women before adjuvant therapy for breast cancer? *Menopause*, *20*(9), 1–8. doi:10.1097/gme.0b013e3182843eff

Bender, C. M., Yasko, J. M., Kirkwood, J. M., Ryan, C., Dunbar-Jacob, J., & Zullo, T. (2000). Cognitive function and quality of life in interferon therapy for melanoma. *Clinical Nursing Research*, *9*(3), 352–363. doi:10.1177/10547730022158537

Berry, D. T. R., Allen, R. S., & Schmitt, F. A. (1991). Rey-Osterrieth complex figure: psychometric characteristics in a geriatric sample. *The Clinical Neuropsychologist*, *5*, 143–153. doi:10.1080/13854049108403298

Bornstein, R. A. (1985). Normative data on selected neuropsychological measures from a nonclinical sample. *Journal of Clinical Psychology*, *41*, 651–669. doi: 10.1002/1097-4679(198509)41:5<651::AID-JCLP2270410511>3.0.CO;2-C

Bousman, C., Szoeke, C., Chen, K., Dennerstein, L., Henderson, V., & Everall, I. (2012). Oestrogen alpha-receptor variant and two-year memory decline in midlife. *Neuropsychobiology*, *66*, 259–265. doi:10.1159/000341879

Boykoff, N., Moieni, M., & Subramanian, S. K. (2009). Confronting chemobrain: an in-depth look at survivors' reports of impact on work, social networks, and health care response. *Journal of Cancer Survivorship : Research and Practice*, *3*(4), 223–232. doi:10.1007/s11764-009-0098-x

Bremer, T. M., Jacquemier, J., Charafe-Jauffret, E., Viens, P., Birnbaum, D., & Linke, S. P. (2009). Prognostic marker profile to assess risk in stage I-III hormone receptor-positive breast cancer patients. *International Journal of Cancer*, *124*(4), 896–904. doi:10.1002/ijc.24001

Brezden, C., Phillips, K., Abdolell, M., Bunston, T., & Tannock, I. F. (2000). Cognitive function in breast cancer patients receiving adjuvant chemotherapy. *Journal of Clinical Oncology*, *18*(14), 2695–2701.

Brinton, R. D., Thompson, R. F., Foy, M. R., Baudry, M., Wang, J., Finch, C. E., … Nilsen, J. (2008). Progesterone receptors: form and function in brain. *Frontiers in Neuroendocrinology*, *29*(2), 313–339. doi:10.1016/j.yfrne.2008.02.001

Britsch, S., Li, L., Kirchhoff, S., Theuring, F., Brinkmann, V., Birchmeier, C., & Riethmacher, D. (1998). The ErbB2 and ErbB3 receptors and their ligand, neuregulin-1, are essential for development of the sympathetic nervous system. *Genes and Development*, *12*(12), 1825–1836. doi:10.1101/gad.12.12.1825

Buyse, M., Loi, S., van't Veer, L., Viale, G., Delorenzi, M., Glas, A. M., … Piccart, M. J. (2006). Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. *Journal of the National Cancer Institute*, *98*(17), 1183–1192. doi:10.1093/jnci/djj329

Callagy, G. M., Webber, M. J., Pharoah, P. D. P., & Caldas, C. (2008). Meta-analysis confirms BCL2 is an independent prognostic marker in breast cancer. *BMC Cancer*, *8*, 153. doi:10.1186/1471-2407-8-153

Capuron, L., Ravaud, A., & Dantzer, R. (2001). Timing and specificity of the cognitive changes induced by interleukin-2 and interferon-alpha treatments in cancer patients. *Psychosomatic Medicine*, *63*(3), 376–386.

Castellon, S. A., Ganz, P. A., Bower, J. E., Petersen, L., Abraham, L., & Greendale, G. A. (2004). Neurocognitive performance in breast cancer survivors exposed to adjuvant chemotherapy and tamoxifen. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 955–969.

Cleeland, C. S. (1989). Measurement of Pain by Subjective Report. In C. R. Chapman & J. D. Loeser (Eds.), *Advances in Pain Research and Therapy* (Vol 12., pp. 391–403). New York: Raven Press.

Cockburn, J., & Smith, P. T. (1993). Correlates of everyday memory among residents of Part III homes. *British Journal of Clinical Psychology*, *32*(Pt 1), 75–77. doi:10.1111/j.2044-8260.1993.tb01029.x

Colleoni, M., Zahrieh, D., Gelber, R., Holmberg, S., Mattsson, J., Rudenstam, C., … Goldhirsch, A. (2005). Site of primary tumor has a prognostic role in operable breast cancer: the International Breast Cancer Study Group experience. *Journal of Clinical Oncology*, *23*(7), 1390–1400. doi:10.1200/JCO.2005.06.052

Dawson, S. J., Makretsov, N., Blows, F. M., Driver, K. E., Provenzano, E., Le Quesne, J., … Pharoah, P. (2010). BCL2 in breast cancer: a favourable prognostic marker across molecular subtypes and independent of adjuvant therapy received. *British Journal of Cancer*, *103*(5), 668-675. doi:10.1038/sj.bjc.6605921

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan (D-KEFS) Executive Function System, Examiners Manual*. San Antonio: The Psychological Corporation.

DNA Genotek Inc. (2012). Laboratory protocol for manual purification of DNA from whole sample. OraSure Technologies, Inc.

Downie, F. P., Mar Fan, H. G., Houédé-Tchen, N., Yi, Q., & Tannock, I. F. (2006). Cognitive function, fatigue, and menopausal symptoms in breast cancer patients receiving adjuvant chemotherapy: evaluation with patient interview after formal assessment. *Psycho-Oncology*, *15*(10), 921–930. doi:10.1002/pon.1035

Dowset, M., Sestak, I., Lopez-Knowles, E., Sidhu, K., Dunbier, A., Cowens, J., … Cuzick, J. (2013). Comparison of PAM50 risk of recurrence score with Oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *Journal of Clinical Oncology*, *31*(22), 2783–3790. doi:10.1200/JCO.2012.46.1558

Edge, S. B., Byrd, D. R., Compton, C. C., Fritz, A. G., Greene, F. L., & Trotti, A. (Eds.). (2010). *AJCC Cancer Staging Manual* (7th ed.). New York, NY: Springer.

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, *19*(5), 403–410. doi:10.1111/j.1365-2559.1991.tb00229.x

Estourgie, S. H., Nieweg, O. E., Olmos, R. A. V., Rutgers, E. J. T., & Kroon, B. B. R. (2004). Lymphatic drainage patterns from the breast. *Annals of Surgery*, *239*(2), 232–237. doi:10.1097/01.sla.0000109156.26378.90

Falleti, M. G., Sanfilippo, A., Maruff, P., Weih, L., & Phillips, K. A. (2005). The nature and severity of cognitive impairment associated with adjuvant chemotherapy in women with breast cancer: a meta-analysis of the current literature. *Brain and Cognition*, *59*(1), 60–70.

doi:10.1016/j.bandc.2005.05.001

Fan, H. G. M., Houédé-Tchen, N., Yi, Q. L., Chemerynsky, I., Downie, F. P., Sabate, K., & Tannock, I. F. (2005). Fatigue, menopausal symptoms, and cognitive function in women after adjuvant chemotherapy for breast cancer: 1- and 2-year follow-up of a prospective controlled study. *Journal of Clinical Oncology*, *23*(31), 8025–8032. doi:10.1200/JCO.2005.01.6550

Fletcher, O., Johnson, N., Orr, N., Hosking, F. J., Gibson, L. J., Walker, K., … Peto, J. (2011). Novel breast cancer susceptibility locus at 9q31.2 : results of a genome-wide association study. *Journal of the National Cancer Institute*, *103*, 425–435. doi:10.1093/jnci/djq563

Genomic Health Inc. (2016). Oncotype DX® Patient Report for Invasive Breast Cancer: Clear Information for Determining Chemotherapy Benefit and Recurrence Risk. Retrieved March 16, 2016, from https://breast-cancer.oncotypedx.com/en-US/Professional-Invasive/Ordering/ReadingTheReports.aspx

Hager, M. H., Morley, S., Bielenberg, D. R., Gao, S., Morello, M., Holcomb, I. N., … Freeman, M. R. (2012). DIAPH3 governs the cellular transition to the amoeboid tumour phenotype. *EMBO Molecular Medicine*, *4*(8), 743–760. doi:10.1002/emmm.201200242

Hazrah, P., Dhir, M., Gupta, S., Deo, V., & Parshad, R. (2009). Prognostic significance of location of the primary tumor in operable breast cancers. *Indian Journal of Cancer*, *46*(2), 139. doi:10.4103/0019-509X.49152

Hein, R., Maranian, M., Hopper, J. L., Kapuscinski, M. K., Southey, M. C., Park, D. J., … Morrison, J. (2012). Comparison of 6q25 breast cancer hits from Asian and European genome wide association studies in the Breast Cancer Association Consortium (BCAC). *PLoS One*, *7*(8), e42380. doi:10.1371/annotation/e5de602c-0ffc-4e6f-a2ed-f79913c2e57c

Henderson, V. W., St John, J. A., Hodis, H. N., Mccleary, C. A., & Stanczyk, F. Z. (2013). Cognition, mood, and physiological concentrations of sex hormones in the early and late postmenopause. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(50), 20290–20295. doi:10.1073/pnas.1312353110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1312353110

Hermelink, K., Untch, M., Lux, M. P., Kreienberg, R., Beck, T., Bauerfeind, I., & Münzel, K. (2007). Cognitive function during neoadjuvant chemotherapy for breast cancer: results of a prospective, multicenter, longitudinal study. *Cancer*, *109*(9), 1905–1913. doi:10.1002/cncr.22610

Hoh, N. Z., Wagner, A. K., Alexander, S. A., Clark, R. B., Beers, S. R., Okonkwo, D. O., … Conley, Y. P. (2010). BCL2 genotypes: functional and neurobehavioral outcomes after severe traumatic brain injury. *Journal of Neurotrauma*, *27*(8), 1413–1427. doi:10.1089/neu.2009.1256

Hong, Y., Chen, X., Li, J., Liu, C., Shen, N., Gong, J., & Chen, W. (2014). Current evidence on the association between rs3757318 of C6orf97 and breast cancer risk : a meta-analysis.

Asian Pacific Journal of Cancer Prevention, *15*(19), 8051–8055. doi:10.7314/APJCP.2014.15.19.8051

Horn, D., Zhou, W., Trevisson, E., Al-Ali, H., Harris, T. K., Salviati, L., & Barrientos, A. (2010). The conserved mitochondrial twin Cx9C protein Cmc2 Is a Cmc1 homologue essential for cytochrome c oxidase biogenesis. *The Journal of Biological Chemistry*, *285*(20), 15088–15099. doi:10.1074/jbc.M110.104786

Hurria, A., Goldfarb, S., Rosen, C., Holland, J., Zuckerman, E., Lachs, M. S., … Hudis, C. (2006). Effect of adjuvant breast cancer chemotherapy on cognitive function from the older patient's perspective. *Breast Cancer Research and Treatment*, *98*(3), 343–348. doi:10.1007/s10549-006-9171-6

Iqbal, N., & Iqbal, N. (2014). Human epidermal growth factor receptor 2 (HER2) in cancers: overexpression and therapeutic implications. *Molecular Biology International*, *2014*, 1–9. doi:10.1155/2014/852748

Iyirhiaro, G. O., Zhang, Y., Estey, C., O'hare, M. J., Safarpour, F., Parsanejad, M., … Park, D. S. (2014). Regulation of ischemic neuronal death by E2F4-p130 protein complexes. *Journal of Biological Chemistry*, *289*(26), 18202–18213. doi:10.1074/jbc.M114.574145

Jansen, C. E., Dodd, M. J., Miaskowski, C. A., Dowling, G. A., & Kramer, J. (2008). Preliminary results of a longitudinal study of changes in cognitive function in breast cancer patients undergoing chemotherapy with doxorubicin and cyclophosphamide. *Psycho-Oncology*, *17*(12), 1189–1195. doi:10.1002/pon.1342

Jasani, B., Novelli, M., Ruschoff, J., & Osamura, R. (2010). HER2 status: what is the difference between breast and gastric cancer? *Connection*, 58–64. Retreived from http://www.dako.com/28830_connection_15_her2_status_what_is_the_difference_between_breast_and_gastric_cancer.pdf

Jenkins, V., Shilling, V., Deutsch, G., Bloomfield, D., Morris, R., Allan, S., … Winstanley, J. (2006). A 3-year prospective study of the effects of adjuvant treatments on cognition in women with early stage breast cancer. *British Journal of Cancer*, *94*(6), 828–834. doi:10.1038/sj.bjc.6603029

Jenkins, V., Shilling, V., Fallowfield, L., Howell, A., & Hutton, S. (2004). Does hormone therapy for the treatment of breast cancer have a detrimental effect on memory and cognition? A pilot study. *Psycho-Oncology*, *13*(1), 61–66. doi:10.1002/pon.709

Jerevall, P. L., Ma, X. J., Li, H., Salunga, R., Kesty, N. C., Erlander, M. G., … Stål, O. (2011). Prognostic utility of HOXB13:IL17BR and molecular grade index in early-stage breast cancer patients from the Stockholm trial. *British Journal of Cancer*, *104*(11), 1762–1769. doi:10.1038/bjc.2011.145

Jiao, L., Zhang, J., Dong, Y., Duan, B., Yu, H., Sheng, H., … Gao, H. (2014). Association between miR-125a rs12976445 and survival in breast cancer patients. *American Journal of Translational Research*, *6*(6), 869–875.

Jim, H. S. L., Donovan, K. A., Small, B. J., Andrykowski, M. A., Munster, P. N., & Jacobsen, P. B. (2009). Cognitive functioning in breast cancer survivors: a controlled comparison. *Cancer*, *115*(8), 1776–1783. doi:10.1002/cncr.24192

Kamakura, T., Akazawa, K., Nomura, Y., Sugimachi, K., & Nose, Y. (1996). Poor prognosis of lower quadrant breast carcinoma. Nishi Nippon Study Group on Adjuvant Chemo-endocrine Therapy for Breast Cancer. *Journal of Surgical Oncology*, *61*(4), 295–9.

Kawamoto, H., Koizumi, H., & Uchikoshi, T. (1997). Expression of the G2-M checkpoint regulators cyclin B1 and cdc2 in nonmalignant and malignant human breast lesions. *American Journal of Pathology*, *150*(1), 15–23.

Kerr, D. A., & Wittliff, J. L. (2011). A five-gene model predicts clinical outcome in ER+/PR+, early-stage breast cancers treated with adjuvant tamoxifen. *Hormones & Cancer*, *2*(5), 261–271. doi:10.1007/s12672-011-0080-8

Kim, J. Y., Sun, Q., Oglesbee, M., & Yoon, S. O. (2003). The role of ErbB2 signaling in the onset of terminal differentiation of oligodendrocytes in vivo. *Journal of Neuroscience*, *23*(13), 5561–5571.

Klein, M. E., Dabbs, D. J., Shuai, Y., Brufsky, A. M., Jankowitz, R., Puhalla, S. L., & Bhargava, R. (2013). Prediction of the Oncotype DX recurrence score: use of pathology-generated equations derived by linear regression analysis. *Modern Pathology*, *26*(5), 658–664. doi:10.1038/modpathol.2013.36

Klove, H. (1963). Clinical neuropsychology. In F. Foster (Ed.), *The Medical Clinics of North America*. New York, NY: Saunders.

Koleck, T. A., & Conley, Y. P. (2016). Identification and prioritization of candidate genes for symptom variability in breast cancer survivors based on disease characteristics at the cellular level. *Breast Cancer: Targets and Thearpy*, *8*, 29–37. doi:http://dx.doi.org/10.2147/BCTT.S88434

Koleck, T., Bender, C., Sereika, S., Ahrendt, G., Jankowitz, R., McGuire, K., … Conley, Y. (2014). Apolipoprotein E genotype and cognitive function in postmenopausal women with early stage breast cancer. *Oncology Nursing Forum*, *41*(6), E313–E325. doi:10.1188/14.ONF.E313-E325

Koleck, T. A., Bender, C. M., Sereika, S. M., Brufsky, A. M., Lembersky, B. C., McAuliffe, P. F., … Conley, Y. P. (2016). Polymorphisms in DNA repair and oxidative stress genes associated with pre-treatment cognitive function in breast cancer survivors: an exploratory study. *SpringerPlus*, *5*, 422. doi:10.1186/s40064-016-2061-4

Kornblum, H. I., Yanni, D. S., Easterday, M. C., & Seroogy, K. B. (2000). Expression of the EGF receptor family members ErbB2, ErbB3, and ErbB4 in germinal zones of the developing brain and in neurosphere cultures containing CNS stem cells. *Developmental Neuroscience*, *22*(1-2), 16–24. doi:10.1159/000017423

Kravitz, H. M., Meyer, P. M., Seeman, T. E., Greendale, G. A., & Sowers, M. R. (2006). Cognitive functioning and sex steroid hormone gene polymorphisms in women at midlife. *The American Journal of Medicine, 119*(9A), 94–102. doi:10.1016/j.amjmed.2006.07.030

Lafayette Clinical Instruments Company. (1989). Lafayette Clinical Repeatable Neuropsychological Battery. Sagamore, IN.

Lee, E., Hsu, C., Haiman, C. A., Razavi, P., Horn-Ross, P. L., van Den Berg, D., … Ursin, G. (2010). Genetic variation in the progesterone receptor gene and risk of endometrial cancer: A haplotype-based approach. *Carcinogenesis*, *31*(8), 1392–1399. doi:10.1093/carcin/bgq113

Lee, K. F., Simon, H., Chen, H., Bates, B., Hung, M. C., & Hauser, C. (1995). Requirement for neuregulin receptor erbB2 in neural and cardiac development. *Nature, 378*, 394–398. doi:10.1038/378394a0

Lejbak, L., Vrbancic, M., & Crossley, M. (2010). Endocrine therapy is associated with low performance on some estrogen-sensitive cognitive tasks in postmenopausal women with breast cancer. *Journal of Clinical and Experimental Neuropsychology*, *32*(8), 836–846. doi:10.1080/13803391003596389

Lengacher, C. A., Reich, R. R., Kip, K. E., Paterson, C. L., Park, H. Y., Ramesar, S., … Park, J. Y. (2015). Moderating effects of genetic polymorphisms on improvements in cognitive impairment in breast cancer survivors participating in a 6-week mindfulness-based stress reduction program. *Biological Research For Nursing*, *17*(4), 393–404. doi:10.1177/1099800415577633

Linke, S. P. (2006). A multimarker model to predict outcome in tamoxifen-treated breast cancer patients. *Clinical Cancer Research*, *12*(4), 1175–1183. doi:10.1158/1078-0432.CCR-05-1562

Liu, D., Biswas, S., & Greene, L. (2004). B-Myb and C-Myb play required roles in neuronal apoptosis evoked by nerve growth factor deprivation and DNA damage. *Journal of Neuroscience, 24*(40), 8720–8725. doi:10.1523/JNEUROSCI.1821-04.2004

Liu, T., Chen, L., Sun, X., Wang, Y., Li, S., Yin, X., … Di, W. (2014). Progesterone receptor PROGINS and +331G/A polymorphisms confer susceptibility to ovarian cancer: a meta-analysis based on 17 studies. *Tumor Biology, 35*(3), 2427–2436. doi:10.1007/s13277-013-1322-x

Lohrisch, B. C., Jackson, J., Jones, A., Mates, D., & Olivotto, I. A. (2000). Relationship between tumor location and relapse in 6,781 women with early invasive breast cancer. *Journal of Clinical Oncology, 18*(15), 2828–2835.

Lowe, C., & Rabbitt, P. (1998). Test\re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: theoretical and practical issues. *Neuropsychologia, 36*(9), 915–923. doi:10.1016/S0028-3932(98)00036-0

Lui, D. X., Nath, N., Chellappan, S. P., & Greene, L. A. (2005). Regulation of neuron survival and death by p130 and associated chromatin modifiers. *Genes and Development*, *19*(6), 719–732. doi:10.1101/gad.1296405

Lyng, M. B., Lænkholm, A. V., Tan, Q., Vach, W., Gravgaard, K. H., Knoop, A., & Ditzel, H. J. (2013). Gene expression signatures that predict outcome of tamoxifen-treated estrogen receptor-positive, high-risk, primary breast cancer patients: a DBCG study. *PLoS One*, *8*(1), e54078. doi:10.1371/journal.pone.0054078

Martelotto, L. G., Ng, C. K. Y., Piscuoglio, S., Weigelt, B., & Reis-Filho, J. S. (2014). Breast cancer intra-tumor heterogeneity. *Breast Cancer Reseatch*, *16*(3), 210. doi:10.1186/bcr3658

Matthews, C. G., Cleeland, C. S., & Hopper, C. L. (1970). Neuropsychological patterns in multiple sclerosis. *Diseases of the Nervous System*, *31*(3), 161–170.

McNair, D., Lorr, M., & Droppleman, L. F. (1992). *EdITS Manual for the Profile of Mood States (POMS)*. San Diego, CA: EdITS/Educational and Industrial Testing Service.

Mehnert, A., Scherwath, A., Schirmer, L., Schleimer, B., Petersen, C., Schulz-Kindermann, F., … Koch, U. (2007). The association between neuropsychological impairment, self-perceived cognitive deficits, fatigue and health related quality of life in breast cancer survivors following standard adjuvant versus high-dose chemotherapy. *Patient Education and Counseling*, *66*(1), 108–118. doi:10.1016/j.pec.2006.11.005

Mandelblatt, J. S., Stern, R. A., Luta, G., McGuckin, M., Clapp, J. D., Hurria, A., … Ahles, T. A. (2014). Cognitive impairment in older patients with breast cancer before systemic therapy: is there an interaction between cancer and comorbidity? *Journal of Clinical Oncology*, *32*(18), 1909-1918. doi: 10.1200/JCO.2013.54.2050

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*(3), 1215. doi:10.1093/nar/16.3.1215

Moasser, M. M. (2007). The oncogene HER2; its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene*, *26*(45), 6469–6487. doi:10.1038/sj.onc.1210477

Moertel, C. G., Reitmeier, R. J., Bolton, C. F., & Schorter, R. G. (1964). Cerebellar ataxia associated with fluorinated pyrimidine therapy. *Cancer Chemotherapy Reports*, *41*, 15–18.

Munir, F., Burrows, J., Yarker, J., Kalawsky, K., & Bains, M. (2010). Women's perceptions of chemotherapy-induced cognitive side affects on work ability: a focus group study. *Journal of Clinical Nursing*, *19*(9-10), 1362–1370. doi:10.1111/j.1365-2702.2009.03006.x

Myers, J. S. (2012). Chemotherapy-related cognitive impairment: the breast cancer experience. *Oncology Nursing Forum*, *39*(1), E31–E40. doi:10.1188/12.ONF.E31-E40

NanoString Technologies Inc. (2013). *NanoString Technologies Receives FDA 510(k) Clearance*

for Prosigna Breast Cancer Prognostic Gene Signature Assay (Vol. 510). Retrieved from http://www.nanostring.com/file/press/PR_Prosigna_FDA_Clearance_130909.pdf

NCBI Resource Coordinators. (2015). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, *43*(D1), D6–D17. doi:10.1093/nar/gku1130

Nelson, H. (1982). *The National Adult Reading Test (NART): Test Manual*. NFER-Nelson.

Ng, T., Teo, S. M., Yeo, H. L., Shwe, M., Gan, Y. X., Cheung, Y. T., … Chan, A. (2015). Brain-derived neurotrophic factor genetic polymorphism (rs6265) is protective against chemotherapy-associated cognitive impairment in patients with early-stage breast cancer. *Neuro-Oncology*, *18*(2), 244-251. doi:10.1093/neuonc/nov162

Ono, M., Ogilvie, J. M., Wilson, J. S., Green, H. J., Chambers, S. K., Ownsworth, T., & Shum, D. H. (2015). A meta-analysis of cognitive impairment and decline associated with adjuvant chemotherapy in women with breast cancer. *Frontiers in Oncology*, *5*(March), 1–19. doi:10.3389/fonc.2015.00059

Osterrieth, P. A. (1944). Le test de copie d'une figure complexe: Contribution a l'etude de la perception et de la memoire. Test of copying a complex figure: contribution to the study of perception and memory. *Archives de Psychologie*, *30*, 206–356.

Owen, A. M., Sahakian, B. J., Semple, J., Polkey, C. E., & Robbins, T. W. (1995). Visuo-spatial short-term recognition memory and learning after temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, *33*(1), 1–24. doi:10.1016/0028-3932(94)00098-A

Paganini-Hill, A., & Clark, L. J. (2000). Preliminary assessment of cognitive function in breast cancer patients treated with tamoxifen. *Breast Cancer Research and Treatment*, *64*(2), 165–176. doi:10.1023/A:1006426132338

Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., … Wolmark, N. (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *The New England Journal of Medicine*, *351*(27), 2817–2826. doi:10.1056/NEJMoa041588

Paik, S., Tang, G., Shak, S., Kim, C., Baker, J., Kim, W., … Wolmark, N. (2006). Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, *24*(23), 3726–34. doi:10.1200/JCO.2005.04.7985

Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T., … Bernard, P. S. (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of Clinical Oncology*, *27*(8), 1160–1167. doi:10.1200/JCO.2008.18.1370

Pavelic, Z. P., Pavelic, L., Lower, E. E., Gapany, M., Gapany, S., Barker, E. A., & Preisler, H. D. (1992). c-myc, c-erbB-2, and Ki-67 expression in normal breast tissue and in invasive and noninvasive breast carcinoma. *Cancer Research*, *52*, 2597–2602.

Polyak, K. (2011). Heterogeneity in breast cancer. *The Journal of Clinical Investigation*, *121*(10), 21–23. doi:10.1172/JCI60534

Powell, J. B., Cripe, L. I., & Dodrill, C. B. (1991). Assessment of brain impairment with the Rey Auditory Verbal Learning Test: a comparison with other neuropsychological measures. *Archives of Clinical Neuropsychology*, *6*(4), 241-249.

Quesnel, C., Savard, J., & Ivers, H. (2009). Cognitive impairments associated with breast cancer treatments: results from a longitudinal study. *Breast Cancer Research and Treatment*, *116*(1), 113–123. doi:10.1007/s10549-008-0114-2

Reed, D. J. (1990). Glutathione: toxicological implications. *Annual Review of Pharmacology and Toxicology*, *30*, 603–631. doi:10.1146/annurev.pa.30.040190.003131

Reitan, R. M. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual & Motor Skills*, *8*, 271–276. doi:10.2466/pms.1958.8.3.271

Reitan, R., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. Tucson, AZ: Neuropsychology Press.

Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. (Les problems). The psychological exaxmination in cases of traumatic encepholopathy. Problems. *Archives de Psychologie*, *28*, 215-285.

Rivenbark, A. G., Connor, S. M. O., & Coleman, W. B. (2013). Molecular and cellular heterogeneity in breast cancer. *The American Journal of Pathology*, *183*(4), 1113–1124. doi:10.1016/j.ajpath.2013.08.002

Robbins, T., James, M., Owen, A., Sahakian, B., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): a factor analytic study of a large sample of normal elderly volunteers. *Dementia and Geriatric Cognitve Disorders*, *5*(5), 266–81. doi:10.1159/000106735

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1997). A neural systems approach to the cognitive psychology of aging using the CANTAB battery. In P Rabbitt (Ed.), *Methodology of Frontal and Executive Function* (pp. 215-238). U. K.: Psychology Press.

Rockwell, L. C., Rowe, E. J., Arnson, K., Jackson, F., Froment, A., Ndumbe, P., … Lorenz, J. G. (2012). Worldwide distribution of allelic variation at the progesterone receptor locus and the incidence of female reproductive cancers. *American Journal of Human Biology*, *24*(1), 42–51. doi:10.1002/ajhb.21233

Romano, A., Delvoux, B., Fischer, D. C., & Groothuis, P. (2007). The PROGINS polymorphism of the human progesterone receptor diminishes the response to progesterone. *Journal of Molecular Endocrinology*, *38*(1-2), 331–350. doi:10.1677/jme.1.02170

Roodi, N., Dupont, W. D., Moore, J. H., & Parl, F. F. (2004). Association of homozygous wild-

type glutathione s-transferase M1 genotype with increased breast cancer risk. *Cancer Research*, *64*, 1233–1236. doi:10.1158/0008-5472.CAN-03-2861

Ryan, C. M., & Williams, T. M. (1993). Effects of insulin-dependent diabetes on learning and memory efficiency in adults. *Journal of Clinical & Experimental Neuropsychology*, *15*(5), 685–700. doi:10.1080/01688639308402589

Ryan, C. M., Williams, T. M., Orchard, T. J., & Finegold, D. N. (1992). Psychomotor slowing is associated with distal symmetrical polyneuropathy in adults with diabetes mellitus. *Diabetes*, *41*(1), 107–113.

Sahakian, B. J., & Coull, J. T. (1993). Tetrahydroaminoacridine (THA) in Alzheimer's disease: an assessment of attentional and mnemonic function using CANTAB. *Acta Neurologica Scandinavica*, *149*, 29–35.

Sala, A. (2005). B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer. *European Journal of Cancer*, *41*(16), 2479–2484. doi:10.1016/j.ejca.2005.08.004

Sarp, S., Fioretta, G., Verkooijen, H. M., Vlastos, G., Rapiti, E., Schubert, H., … Bouchardy, C. (2007). Tumor location of the lower-inner quadrant is associated with an impaired survival for women with early-stage breast cancer. *Annals of Surgical Oncology*, *14*(3), 1031–1039. doi:10.1245/s10434-006-9231-5

Schagen, S. B., van Dam, F. S., Muller, M. J., Boogerd, W., Lindeboom, J., & Bruning, P. F. (1999). Cognitive deficits after postoperative adjuvant chemotherapy for breast carcinoma. *Cancer*, *85*(3), 640–650. doi:10.1002/(SICI)1097-0142(19990201)85:3<640::AID-CNCR 14>3.0.CO;2-G

Schilder, C. M., Eggens, P. C., Seynaeve, C., Linn, S. C., Boogerd, W., Gundy, C. M., … Schagen, S. B. (2009). Neuropsychological functioning in postmenopausal breast cancer patients treated with tamoxifen or exemestane after AC-chemotherapy: cross-sectional findings from the neuropsychological TEAM-side study. *Acta Oncologica*, *48*(1), 76–85. doi:10.1080/02841860802314738

Schilder, C. M., Seynaeve, C., Beex, L. V, Boogerd, W., Linn, S. C., Gundy, C. M., … Schagen, S. B. (2010). Effects of tamoxifen and exemestane on cognitive functioning of postmenopausal patients with breast cancer: results from the neuropsychological side study of the tamoxifen and exemestane adjuvant multinational trial. *Journal of Clinical Oncology*, *28*(8), 1294–1300. doi:10.1200/JCO.2008.21.3553

Schwab, R., Bussolari, R., Corvetta, D., Chayka, O., Santilli, G., Kwok, J. M., … Sala, A. (2008). Isolation and functional assessment of common, polymorphic variants of the B-MYB proto-oncogene associated with a reduced cancer risk. *Oncogene*, *27*(20), 2929–2933. doi:10.1038/sj.onc.1210947

Seigers, R., Pourtau, L., Schagen, S. B., van Dam, F. S., Koolhaas, J. M., Konsman, J. P., & Buwalda, B. (2010). Inhibition of hippocampal cell proliferation by methotrexate in rats is not potentiated by the presence of a tumor. *Brain Research Bulletin*, *81*(4-5), 472–476.

doi:10.1016/j.brainresbull.2009.10.006

Shilling, V., & Jenkins, V. (2007). Self-reported cognitive problems in women receiving adjuvant therapy for breast cancer. *European Journal of Oncology Nursing*, *11*(1), 6–15. doi:10.1016/j.ejon.2006.02.005

Shilling, V., Jenkins, V., Fallowfield, L., & Howell, T. (2003). The effects of hormone therapy on cognition in breast cancer. *The Journal of Steroid Biochemistry and Molecular Biology*, *86*(3-5), 405–412. doi:10.1016/j.jsbmb.2003.07.001

Silva, A. R., Santos, A. C., Farfel, J. M., Grinberg, L. T., Ferretti, R. E., Campos, A. H., … Brentani, H. (2014). Repair of oxidative DNA damage, cell-cycle regulation and neuronal death may influence the clinical manifestation of Alzheimer's disease. *PloS One*, *9*(6), e99897. doi:10.1371/journal.pone.0099897

Silvestre-Roig, C., Fernández, P., Mansego, M. L., van Tiel, C. M., Viana, R., Anselmi, C. V., … Andrés, V. (2014). Genetic variants in CCNB1 associated with differential gene transcription and risk of coronary in-stent restenosis. *Circulation. Cardiovascular Genetics*, *7*(1), 59–70. doi:10.1161/CIRCGENETICS.113.000305

Sloan, C. D., Shen, L., West, J. D., Wishart, H. A., Flashman, L. A., Rabin, L. A., … Saykin, A. J. (2010). Genetic pathway-based hierarchical clustering analysis of older adults with cognitive complaints and amnestic mild cognitive impairment using clinical and neuroimaging phenotypes. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *153*(5), 1060–1069. doi:10.1002/ajmg.b.31078

Small, B. J., Rawson, K. S., Walsh, E., Jim, H. S. L., Hughes, T. F., Iser, L., … Jacobsen, P. B. (2011). Catechol-o-methyltransferase genotype modulates cancer treatment-related cognitive deficits in breast cancer survivors. *Cancer*, *117*(7), 1369–1376. doi:10.1002/cncr.25685

Snow, W. G., Tierney, M. C., Zorzitto, M. L., Fisher, R. H., & Reid, D. W. (1989). WAIS-R test-retest reliability in a normal elderly sample. *Journal of Clinical & Experimental Neuropsychology*, *11*(4), 423–428. doi:10.1080/01688638908400903

Sohn, V. Y., Arthurs, Z. M., Sebesta, J. A., & Brown, T. A. (2008). Primary tumor location impacts breast cancer survival. *American Journal of Surgery*, *195*(5), 641–644. doi:10.1016/j.amjsurg.2007.12.039

Stark, B. A., Hulka, B. S., Joens, S., Novotny, D., Thor, A. D., Wold, L. E., … Conway, K. (2000). HER-2/neu amplification in benign breast disease and the risk of subsequent breast cancer. *Journal of Clinical Oncology*, *18*(2), 267–274.

Stenzig, J., Schweikert, A., Piasecki, A., Höppner, G., Eschenhagen, T., & Rau, T. (2012). Progesterone receptor variants associated with the PROGINS haplotype exhibit functional properties similar to those of wild-type progesterone receptor. *Pharmacogenetics and Genomics*, *22*(8), 629–641. doi:10.1097/FPC.0b013e3283558256

Stevens, K. N., Vachon, C. M., Lee, A. M., Slager, S., Lesnick, T., Olswold, C., … Hein, R. (2011). Common breast cancer susceptibility loci are associated with triple-negative breast cancer. *Cancer Research*, *71*(19), 6240–6250. doi:10.1158/0008-5472.CAN-11-1266

Stewart, A., Bielajew, C., Collins, B., Parkinson, M., & Tomiak, E. (2006). A meta-analysis of the neuropsychological effects of adjuvant chemotherapy treatment in women treated for breast cancer. *The Clinical Neuropsychologisteuropsychologist*, *20*(1), 76–89. doi:10.1080/138540491005875

Su, Y., Jiang, Y., Sun, S., Yin, H., Shan, M., Tao, W., … Pang, D. (2015). Effects of HER2 genetic polymorphisms on its protein expression in breast cancer. *Cancer Epidemiology*, *39*(6), 1123–1127. doi:10.1016/j.canep.2015.08.011

Sundermann, E., Maki, P., & Bishop, J. (2011). A review of estrogen receptor α gene (ESR1) polymorphisms, mood, and cognition. *Menopause*, *17*(4), 874–886. doi:10.1097/gme.0b013e3181df4a19.A

Tager, F. A., McKinley, P. S., Schnabel, F. R., El-Tamer, M., Cheung, Y. K., Fang, Y., … Hershman, D. L. (2010). The cognitive effects of chemotherapy in post-menopausal breast cancer patients: a controlled longitudinal study. *Breast Cancer Research and Treatment*, *123*(1), 25–34. doi:10.1007/s10549-009-0606-8.

Tavassoli, F. A., & Devilee, P. (Eds.). (2003). *World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of the Breast and Female Genital Organs.* Lyon, France: IARC Press.

Tchen, N., Juffs, H. G., Downie, F. P., Yi, Q. L., Hu, H., Chemerynsky, I., … Tannock, I. F. (2003). Cognitive function, fatigue, and menopausal symptoms in women receiving adjuvant chemotherapy for breast cancer. *Journal of Clinical Oncology*, *21*(22), 4175–4183. doi:10.1200/JCO.2003.01.119

Thompson, M., Lauderdale, S., Webster, M. J., Chong, V. Z., McClintock, B., Saunders, R., & Weickert, C. S. (2007). Widespread expression of ErbB2, ErbB3 and ErbB4 in non-human primate brain. *Brain Research*, *1139*(1), 95–109. doi:10.1016/j.brainres.2006.11.047

Tutt, A., Wang, A., Rowland, C., Gillett, C., Lau, K., Chew, K., … Waldman, F. (2008). Risk estimation of distant metastasis in node-negative, estrogen receptor-positive breast cancer patients using an RT-PCR based prognostic expression signature. *BMC Cancer*, *8*, 339. doi:10.1186/1471-2407-8-339

Urruticoechea, A., Smith, I. E., & Dowsett, M. (2005). Proliferation marker Ki-67 in early breast cancer. *Jornal of Clinical Oncology*, *23*(28), 7212–7220. doi:10.1200/JCO.2005.07.501

van Dam, F. S. A. M., Schagen, S. B., Muller, M. J., Boogerd, W., Wall, E. V. D., Fortuyn, M. E. D., & Rodenhuis, S. (1998). Impairment of cognitive function in women receiving adjuvant treatment for high-risk breast cancer: high-dose versus standard-dose chemotherapy. *Journal of the National Cancer Institute*, *90*(3), 210–218. doi:10.1093/jnci/90.3.210

van't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., … Friend, S. H. (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, *415*(6871), 530–536. doi:10.1038/415530a

Vardy, J., Wefel, J. S., Ahles, T., Tannock, I. F., & Schagen, S. B. (2008). Cancer and cancer-therapy related cognitive dysfunction: an international perspective from the Venice cognitive workshop. *Annals of Oncology*, *19*(4), 623–629. doi:10.1093/annonc/mdm500

Veeraraghavan, J., Tan, Y., Cao, X., Kim, J., Chamness, G. C., Maiti, S. N., … Wang, X. (2015). Recurrent ESR1-CCDC170 rearrangements in an aggressive subset of estrogen-receptor positive breast cancers. *Nature Communications*, *7*(5), 4577. doi:10.1038/ncomms 5577.Recurrent

Von Ah, D., Habermann, B., Carpenter, J. S., & Schneider, B. L. (2012). Impact of perceived cognitive impairment in breast cancer survivors. *European Journal of Oncology Nursing*, *17*(2), 236-241. doi:10.1016/j.ejon.2012.06.002

Von Ah, D., Harvison, K. W., Monahan, P. O., Moser, L. R., Zhao, Q., Carpenter, J. S., … Unverzagt, F. W. (2009). Cognitive function in breast cancer survivors compared to healthy age- and education-matched women. *The Clinical Neuropsychologist*, *23*(4), 661–674. doi:10.1080/13854040802541439

Vorstman, A., van Daalen, E., Jalali, G., Schmidt, E., Pasterkamp, R., de Jonge, M., … Ophoff, R. (2011). A double hit implicates DIAPH3 as an autism risk gene. *Molecular Psychiatry*, *72*(2), 181–204. doi:10.1038/nature13314.A

Voytko, M. L., Murray, R., & Higgs, C. J. (2009). Executive function and attention are preserved in older surgically menopausal monkeys receiving estrogen or estrogen plus progesterone. *Journal of Neuroscience*, *29*(33), 10362–10370. doi:10.1523/JNEUROSCI.1591-09.2009

Wang, V., Chuang, T. C., Kao, M. C., Shan, D. E., Soong, B. W., & Shieh, T. M. (2013). Polymorphic Ala-allele carriers at residue 1170 of HER2 associated with Parkinson's disease. *Journal of the Neurological Sciences*, *325*(1-2), 115–119. doi:10.1016/ j.jns.2012.12.017

Wechsler, D. (1981). *The Wechsler Adult Intelligence Scale-Revised manual*. New York, NY: Psychological Corporation.

Wechsler, D. (1998). *The Wechsler Memory Scale-Revised manual*. San Antonio, TX: Psychological Corporation.

Wefel, J., Kesler, S., Noll, K., & Schagen, S. (2015). Clinical characteristics, pathophysiology, and management of noncentral nervous system cancer-related cognitive impairment in adults. *CA: A Cancer Journal for Clinicians*, *65*(2), 123–138. doi:10.3322/caac.21258.

Wefel, J. S., Lenzi, R., Theriault, R., Buzdar, A. U., Cruickshank, S., & Meyers, C. A. (2004). "Chemobrain" in breast carcinoma?: a prologue. *Cancer*, *101*(3), 466–475. doi:10.1002/cncr.20393

Wefel, J. S., Saleeba, A. K., Buzdar, A. U., & Meyers, C. A. (2010). Acute and late onset cognitive dysfunction associated with chemotherapy in women with breast cancer. *Cancer*, *116*(14), 3348–3356. doi:10.1002/cncr.25098

Wefel, J. S., & Schagen, S. B. (2012). Chemotherapy-related cognitive dysfunction. *Current Neurology and Neuroscience Reports*, *12*(3), 267–275. doi:10.1007/s11910-012-0264-9

Wesoła, M., & Jeleń, M. (2015). A comparison of IHC and FISH cytogenetic methods in the evaluation of HER2 status in breast cancer. *Advances in Clinical and Experimental Medicine*, *24*, 899–904. doi:10.17219/acem/27923

Wieneke, M. H., & Dienst, E. R. (1995). Neuropsychological assessment of cognitive functioning following chemotherapy for breast cancer. *Psycho-Oncology*, *4*(2), 61–66. doi:10.1002/pon.2960040108

Wilson, B., Cockburn, J., Baddeley, A., & Hiorns, R. (1989). The development and validation of a test battery for detecting and monitoring everyday memory problems. *Journal of Clinical & Experimental Neuropsychology*, *11*(6), 855–870. doi:10.1080/01688638908400940

Winters, Z. E., Hunt, N. C., Bradburn, M. J., Royds, J. A., Turley, H., Harris, A. L., & Norbury, C. J. (2001). Subcellular localisation of cyclin B, Cdc2 and p21WAF1/CIP1 in breast cancer: association with prognosis. *European Journal of Cancer*, *37*, 2405–2412. doi:10.1016/S0959-8049(01)00327-6

Wolff, A. C., Hammond, M. E. H., Schwartz, J. N., Hagerty, K. L., Allred, D. C., Cote, R. J., … Hayes, D. F. (2007). American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *Archives of Pathology & Laboratory Medicine*, *131*(1), 18–43. doi:10.1043/1543-2165(2007)131[18:ASOCCO]2.0.CO;2

Wu, S., Zhou, J., Ren, Y., Sun, J., Li, F., Lin, Q., … He, Z. (2014). Tumor location is a prognostic factor for survival of Chinese women with T1-2N0M0 breast cancer. *International Journal of Surgery*, *15*(5), 394–398. doi:10.1016/j.ijsu.2014.03.011

Yaffe, K., Lindquist, K., Sen, S., Cauley, J., Ferrell, R., Penninx, B., … Cummings, S. R. (2009). Estrogen receptor genotype and risk of cognitive impairment in elders: Findings from the Health ABC study. *Neurobiology of Aging*, *30*, 607–614. doi:10.1016/j.neurobiolaging.2007.08.003

Yaffe, K., Lui, L., Grady, D., Stone, K., & Morin, P. (2002). Estrogen receptor 1 polymorphisms and risk of cognitive impairment in older women. *Biological Psychiatry*, *51*, 677–682. doi:10.1016/S0006-3223(01)01289-6

Yamamoto-Ibusuki, M., Yamamoto, Y., Fujiwara, S., Sueta, A., Yamamoto, S., Hayashi, M., … Iwase, H. (2015). C6ORF97-ESR1 breast cancer susceptibility locus : influence on progression and survival in breast cancer patients. *European Journal of Human Genetics*, *23*(7), 949–956. doi:10.1038/ejhg.2014.219

Youngjohn, J. R., Larrabee, G. J., & Crook, T. H. (1992). Test-retest reliability of computerized, everyday memory measures and traditional memory tests. *The Clinical Neuropsychologist*, *6*(3), 276–286. doi:10.1080/13854049208404129

Yu, K. D., Di, G. H., Fan, L., Wu, J., Hu, Z., Shen, Z. Z., … Shao, Z. M. (2009). A functional polymorphism in the promoter region of GSTM1 implies a complex role for GSTM1 in breast cancer. *The FASEB Journal*, *23*(7), 2274–2287. doi:10.1096/fj.08-124073

Zwart, W., Terra, H., Linn, S. C., & Schagen, S. B. (2015). Cognitive effects of endocrine therapy for breast cancer: keep calm and carry on? *Nature Reviews Clinical Oncology*, *12*(10), 597–606. doi:10.1038/nrclinonc.2015.124