FINE MAPPING REGIONS OF EXCESS HOMOZYGOSITY-BY-DESCENT IN PALAU

by

Cara S. Dresbold

BS, Pennsylvania State University, 2010

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH
GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Cara S. Dresbold

It was defended on

March 6, 2013

and approved by

Robin E. Grubs, PhD, CGC
Assistant Professor
Department of Human Genetics
Co-Director, Genetic Counseling Program
Graduate School of Public Health
University of Pittsburgh

Nadine Melhem, PhD
Assistant Professor of Psychiatry
Department of Psychiatry
School of Medicine
University of Pittsburgh

Robert E. Ferrell, PhD
Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Thesis Advisor:
Bernie Devlin, PhD
Professor
Department of Human Genetics
Graduate School of Public Health
Professor, Department of Psychiatry
School of Medicine
University of Pittsburgh

FINE MAPPING REGIONS OF EXCESS HOMOZYGOSITY-BY-DESCENT IN PALAU

Cara S. Dresbold, MS

University of Pittsburgh, 2013

ABSTRACT

Schizophrenia is a devastating psychiatric disorder that affects approximately 1% of the general population. Its complexity and enigmatic etiology place an economic and social burden on public health resources, while causing emotional distress for patients and their families. Twin, family, and adoption studies have shown evidence for a genetic component, but its complex pattern of inheritance has complicated the efforts for identifying susceptibility genes. The discovery of disease causing genes has public health relevance given that it could lead the path to pre-symptomatic testing, preventative care, and a better prognosis. Recent studies have shown that inbreeding, through consanguineous marriages, is associated with increased risk for schizophrenia. In the genetically isolated island of Palau, the prevalence of schizophrenia is more than doubled at 2.7%. Here, we examine the role of excess homozygosity and regions of homozygosity-by-descent on schizophrenia risk in Palau. Palauans practice exogamy, but the likely limited size of the founding population, combined with recent population bottlenecks, ensures that homozygous segments of the genome, shared identical-by-descent from a common ancestor, occur regularly in Palauan genomes. Five gene regions, *RAB3GAP2*, *MARK1*, *DISC1*, *FAM184A*, and *IGSF9B* were found to overlap segments that were commonly shared among subjects who were homozygous-by-descent. Cases and controls that were homozygous for these regions were counted and compared, with heightened representation from cases. Further SNP

haplotype and pedigree analysis, however, determined that none of the regions showed evidence of a highly penetrance recessive risk locus for schizophrenia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank Bernie Devlin, PhD, and Nadine Melhem, PhD for their leadership and supervision throughout the duration of the project. Thank you to committee members Robin Grubs, PhD, CGC, and Robert Ferrell, PhD for their continuous guidance and support. Thank you to Shawn Wood and Lambertus Klei for their assistance in the computational analysis portion of the thesis. The project would not have been a success without all of their collaborated input and participation.

1.0 INTRODUCTION

Schizophrenia is a devastating complex psychiatric disease that affects approximately 1% of the general population. In the Palauan population, the occurrence of the disease is more than doubled (2.7%). It is possible this slight elevation is simply stochastic, but an intriguing possibility is that it has a biological origin. Palau, like many Oceanic populations has experienced a series of bottleneck events throughout its history. These, together with what was likely a small population size at founding, increases the genetic relationships among mating partners even in the face of the regular outbreeding practiced by Palauans. Given Palau's geographical isolation and reduced gene flow, it could be an ideal candidate for gene mapping studies aimed to identify genes involved in conferring risk for schizophrenia. Perhaps cryptic inbreeding and its resultant impact on expression of recessive alleles could account for the increased rate of schizophrenia in Palau. Lu et al. (unpublished data) found that there is an association between diagnosis of schizophrenia in Palau and long chromosomal segments that are homozygous-by-descent. It is possible that homozygosity is associated with risk for schizophrenia, but the mechanism generating this association is not obvious, and specific gene regions harboring highly penetrant recessive alleles have yet to be identified.

2.0 BACKGROUND AND SIGNIFICANCE

2.1 SCHIZOPHRENIA

2.1.1 Epidemiology

Schizophrenia is a complex psychiatric illness that presents similarly in all ethnic groups across the globe. It has consistently been shown to affect 1% of the world's population, despite presenting among different cultures and in different environments (3, 6). Approximately 1 in 10,000 individuals worldwide are newly diagnosed with schizophrenia each year (1). The onset of symptoms typically occurs between the ages of 16 and 30 years old, and although men and women can both be affected with the disease, men tend to experience symptoms at earlier ages (4). On average, the age of onset for men is during their early to mid 20's, whereas women typically begin developing their first symptoms in their late 20's (1). Studies have shown that prognosis of schizophrenia may be associated with a variety of factors. These factors include early age of onset, being male, having a lower level of education, and having pre-symptomatic adjustment issues (1). Females who show signs of the condition later in life tend to have a less severe prognosis.

2.1.2 Clinical Features of Schizophrenia

Schizophrenia is identified by a constellation of many cognitive and emotional symptoms. These symptoms can affect an individual's logical thinking, speech, behavior, perception, motivations, and attention (1). The clinical manifestations of schizophrenia are categorized into two subsets: positive and negative symptoms.

2.1.2.1 Positive Symptoms

Positive symptoms represent features of schizophrenia that were not present in the individual prior to the onset of symptoms, and have an excess or distortion of normal functions (1). Potential examples of this subset of symptoms include delusions, hallucinations, disorganized speech, catatonic behavior, or inappropriate affect (2, 5). The DSM-IV defines delusions as "erroneous beliefs that usually involve a misinterpretation of perceptions or experiences" (1). The most common form of delusion is persecutory delusion in which the patient believes that harm will intentionally be placed on him or her by others. Referential delusions are egocentric beliefs that external references, such as newspaper headings or song lyrics, are specifically aimed at the patient. Delusions are considered bizarre if they are completely implausible and are unable to be derived from reality (2). These thoughts are strongly retained by the individual despite evidence of the contrary.

Hallucinations are sensory experiences that seem present to the affected individual, but are not justified by what is occurring in reality. Hallucinations can appear as auditory, visual, olfactory, gustatory, or tactile sensations (1). Auditory hallucinations are the most common form, which often include hearing voices in one's head. This specific finding by itself is enough to confirm a diagnosis of schizophrenia.

Disorganization in thinking, speech, or behavior can present in a variety of ways. Those whose thinking is non-linear and unable to stay on-track often express this through their speech (1). Conversations consist of loose associations and may even be incomprehensible to the listener. Disorganized behavior can range from unusual agitation to complete impairment. Some individuals have shown the inability to maintain daily activities such as getting dressed or practicing hygiene. Other individuals appear to be completely unaware of their environment. Catatonic motor behaviors can occur when an individual loses all reactivity to their surroundings (1). A variety of motions can represent these behaviors, including remaining in a rigid posture even when intentionally moved, and becoming over stimulated and excited without reason (1). Although these symptoms are not necessarily specific to schizophrenia, their combined presence in an individual can aid in the attempt to determine a diagnosis.

2.1.2.2 Negative Symptoms

Negative symptoms appear when there is a loss or disruption in a behavior that was present prior to disease onset. Patients may experience a range of emotions or moods including alogia, affective flattening, avolition, or anhedonia. (5). Alogia includes loss of speech production, reflected in decreased verbal response and fluency (1). Affective flattening is reflected in an individual's body language, and can include poor eye contact, immobility, and reduced facial expression (1). Avolition is characterized by the loss of motivation. It can often be seen in an individual who has stopped pursuing goal-oriented activities and who has lost interest in work or social activities (1). Anhedonia is not included in the definition of schizophrenia, but it is considered an associative feature. It is characterized by the inability to experience pleasure in activities and situations that were at one time pleasurable to the individual.

2.1.3 Diagnosis

A patient is diagnosed with schizophrenia based on specific criteria that are outlined in the 1994 edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (1, 2). The requirements for diagnosis include specific symptomatic findings, duration, and exclusion of comparable psychiatric disorders (2). Schizophrenia is diagnosed when there is a manifestation of at least two symptoms lasting the duration of six months, including one positive symptom lasting a minimum of one month (1). The presence of only one symptom may be sufficient for making a diagnosis if it represents a bizarre delusion or an auditory hallucination that consists of hearing non-existing voices (1).

Symptoms associated with schizophrenia will often disturb levels of social functioning, affecting one's ability to maintain a job, interact appropriately with others, and partake in self-care practices. If the outcome of the symptoms do not affect the ability to function in society, it is unlikely to be diagnosed as a disorder. Similar symptoms are found in other psychiatric disorders including mood and schizo-affective disorders (2). These conditions, as well as other potential medical causes for the abnormal behaviors, must be excluded before a proper diagnosis of schizophrenia can be finalized. These disorders are likely to be ruled out if a patient does not show signs more highly associated with these conditions including major depression, manic, or mixed episodes (1). In addition, the use of certain drugs and medications that may be contributing to the development of symptoms must also be investigated and ruled out prior to a diagnosis (1).

2.1.4 Evidence for a Genetic Component

Similar to other psychiatric disorders, identifying genes for schizophrenia has been complicated by its complexity. It is possible that its intricacies are a result of a range of factors including phenocopies, genetic heterogeneity, variable expression, polygenic inheritance, as well as environmental components (13). Despite the challenges, the evidence that schizophrenia is influenced by a genetic component continues to accumulate. Family, twin, and adoption studies provide evidence for a genetic component in schizophrenia.

When genetics play a role in any condition, it is expected that the condition will occur more often among those related to affected individuals than those in the general population. There is an increase in the incidence of schizophrenia in family members of affected individuals, as well as a correlation in the level of risk with the degree of relation among members (6). As the portion of the genome that is shared identical-by-descent (IBD) increases, so does the risk for developing schizophrenia (7). Table 1 reports the values, reported by major European studies, for risk in reference to familial relationships (31). Although these numbers vary by study, the direction of the pattern remains the same. In general, the risk for schizophrenia in first degree relatives is 5-10 times greater compared to the risk in the general population (1, 7). However, family studies cannot differentiate the environmental factors that are affecting risks as well.

Twin studies provide evidence for the role of genetics in schizophrenia. A monozygotic twin of an affected individual is four times more likely to develop schizophrenia compared to a dizygotic twin (7). Both sets of twins share both pre and postnatal environments at the same level, yet monozygotic twins share complete genetic information, whereas dizygotic twins share, on average, only 50% of their genetic variation. Concordance rates for schizophrenia among identical twins have been reported to be between 50-57.7% and 12.8-15% for fraternal twins (6,

7). Even in the situation where there is only one member of monozygotic twins affected with schizophrenia, the other twin is more likely to present with other disorders of psychosis or mental health issues (7). Based on twin studies, it is estimated that the heritability for schizophrenia is approximately 60-90% (3, 7, 11). Narrow-sense heritability is defined as "the proportion of phenotypic variance in a population attributable to additive genetic factors" (26).

Adoption studies have the unique quality of being able to focus on biological risk for schizophrenia while removing the impact of postnatal environmental factors (1). Studies have shown that children adopted away from schizophrenic mothers are at a higher risk of developing the disease than children who were adopted away from unaffected mothers (7.7-11% compared to 0%) (7). This suggested that it is the biological, not the adoptive, relationship that is associated with familial transmission. Other studies went further to assess the risk to adopted paternal half siblings who shared biological fathers affected by the disease. Given that these half siblings did not share the same mother, they did not share the same prenatal environment. Since they were adopted, they also did not share the same postnatal environment. They only shared the genetic information passed from their mutual father. The results of these studies showed that adopted out half siblings who had shared a schizophrenic father were at greater risk of developing schizophrenia than adopted out half siblings who did not have a schizophrenic father, with 14 half-siblings of affected cases developing schizophrenia compared to only two half-siblings of controls (32). This information continues to build on the idea that there is a genetic component involved in the risk for schizophrenia.

Table 1. Familial relationships and risk for schizophrenia (31)

Relation	*Risk for schizophrenia*
Parents	4.4%
Sibling (neither parent schizophrenic)	8.2%
Sibling (one parent schizophrenic)	13.8%
Fraternal twins of opposite sex	5.6%
Fraternal twins of same sex	12.0%
Identical twins	57.7%
Children	12.3%
Children (both parents schizophrenic)	36.6%
Uncles/Aunts	2.0%
Nephews/Nieces	2.2%
Grandchildren	2.8%
Half siblings	3.2%
First cousins	2.9%
General Population	0.86%

2.2 PALAU

2.2.1 Ethnography

The Republic of Palau is an archipelago made up of over 200 islands that cover a 125 mile stretch (16). These islands are located in the westernmost part of Micronesia, far from any other major land mass. Settlement in the region began approximately 2,000 years ago, when migrants were believed to have come over from the Philippines or Indonesia, based on blood group clustering, linguistic analysis, and ethnographic studies. (15). Archeological studies have shown that the population has grown primarily in isolation from the point of the original settlement up until the last two centuries (9, 15). Throughout the history of Palau, a series of bottleneck events have taken place, resulting in the manipulation of its population size and dynamic. During World

War II, the population was reduced to only 3,500 natives. Today, it has grown to a population size of approximately 21,000 (16).

Ethnographic characteristics of Palau led to geographic and genetic isolation. Through the generations, the cultural dynamic of Palau was made up of distinguished blood-based clans (9). Clans consisted of closely related families that were derived from a common female founder. The configuration of these clans has been continuously preserved by the maintenance of accurate records from excellent historians within the groups (15). Despite practicing exogamy, genetic studies suggest that there are low levels of gene flow among the Palauans, likely due to a small founder population, several bottlenecks, and longstanding geographical isolation (27, 39). Limitations of the range of genetic heterogeneity have an effect on the genetic makeup of living Palauans, and may be contributing to increase risk of complex diseases among the population.

2.2.2 Founder Effects

Founder effects arise when the initial population of a geographic area is small and demographic growth occurs with limited migration (8). The small size of the population rapidly changes the frequency of any given allele over few generations, relative to a population with a larger effective population size. If one or more individuals in the founding population carried a genetic mutation associated with disease, the frequency of the mutation could increase rapidly in the population with each subsequent generation; alternatively, it could disappear from the population very rapidly. For the former to occur, reproductive abilities of carriers would have to remain intact.

If the frequency of the mutation was truly increased, there would be a greater chance that matings between members of the population would produce offspring that were homozygous for

that specific mutation and it could lead to a greater frequency of disease. It has also been suggested that diseases that follow multifactorial inheritance may also be at higher frequencies in populations that have undergone founder effects.

2.2.3 Schizophrenia in Palau

In Palau, the prevalence of schizophrenia estimated to be 2.7%, which is more than double that of the worldwide estimates (9). This finding is not unique to Palau, in that many isolated island nations report increased frequency of schizophrenia, including 2.4% in North Sweden, and 3.2% in Finland (15). In Palau, the highest rates of schizophrenia occur among natives living in the western region of the islands, and the lowest prevalence occurs in the east. It is predicted that this may reflect the lack of genetic flow through the islands because of the mating practices set by the clans. In addition, the breeding of close relatives have been shown to be associated with an increased risk of several other complex diseases like schizophrenia, including diabetes, cancer, and heart disease (17). In several Middle Eastern countries, consanguineous marriages between first and second cousins are more likely to result in offspring with schizophrenia than in marriages between distantly related partners (17).

Several studies have used the Palauan population in an attempt to discover genes and loci related to risk for schizophrenia. One study examined copy number variants (CNVs) that conferred risk for schizophrenia in subjects of European ancestry in Palauan subjects. Of the CNVs known or thought to convey risk for schizophrenia in other populations (18, 33), several were detected in the Palau sample, including *1q21.1*, *15q11.2*, *16p13.2*, and *Xp21.3* (18). It is believed that Palau's small founder population, geographic isolation, history of bottleneck

events, and high prevalence of schizophrenia compared to the general population, makes the population an ideal candidate in the continuing search for the genetic etiology of schizophrenia.

2.3 RISK LOCI IDENTIFIED OUTSIDE PALAU

Many studies have investigated the molecular etiology of schizophrenia and related psychiatric disorders through genome-wide association studies in various populations. Mega-analysis of genome-wide association studies found seven regions that were associated with increased risk for schizophrenia in individuals of European ancestry. These regions, including 1p21.3, 2q32.3, 8p23.2, 8q21.3, 10q24.32-q24.33, 6p21.32-p22.1, and 18q21.2, were by using logistic regression of imputed "dosages" (fractional allocations of genotypes based on imputed genotype probabilities) against diagnostic phenotype (case/control status) (43). Region 6p21.32-p22.1 maps to the major histocompatibility complex (MHC) which has previously been implicated in schizophrenia (43). The strongest signal that had not been described previously was found at region 1p21.3 ($p=1.6x10^{-11}$). This region contains the primary transcript for *MIR137*. This gene is known as a regulator for adult neurogenesis and neuronal maturation (43). This finding paired with imaging studies that identified targets significant for association with schizophrenia suggest that *MIR137*-mediated dysregulation may be an etiologic mechanism in schizophrenia.

By cross-disorder analyses, or meta-analyses involving schizophrenia, as well as other psychiatric conditions including autism spectrum disorders, attention deficit-hyperactivity disorder, bipolar disorder, and major depressive disorder, genome-wide association studies have identified additional regions that are associated with risk (40). These regions include

chromosome 2q32.1 which overlaps with gene *ZNF804A*, an intronic SNP within gene *ITIH3* on chromosome 3p21, an intron of gene *AS3MT* on chromosome 10q24, a region of chromosome 12 within *CACNA1C*, and an intron of *CACNB2* on chromosome 10 (40, 41, 43). Both *CACNA1C* and *CACNB2* are genes highly expressed in the brain that encode L-type voltage-gated calcium channel subunits (40). Previous studies have identified *CACNA1C* as a susceptibility gene for schizophrenia, bipolar disorder, and major depressive disorder. Gain of function mutations in this gene are associated with a developmental disorder known as Timothy syndrome, which has autism as a major presenting feature (40). *CACNB2* is an auxiliary voltage-gated calcium channel that is known to interact with *CACNA1C* to increase its function. Variants in this gene were found to be a main signal in bipolar disorder in affected individuals from the Han Chinese population (40). Significant association between these genes and the development of psychiatric disorders may suggest that voltage-gated calcium channel activity could be an important biological process in psychiatric disorders, including schizophrenia (40). Although not exclusive to schizophrenia, these findings could play a part in determining the molecular etiology that leads to a variety psychiatric diseases.

Loci conveying risk for schizophrenia have also been found through the presence of rare copy number variants (CNV) at a particular region, especially those that cause the disruption of relevant genes. Many CNVs have been identified in samples of European ancestry that are believed to be associated with an increased risk for psychiatric conditions. For instance, deletions in the velocardiofacial syndrome region located on chromosome 22q11.2, are associated with a 20 fold increased risk for psychiatric disorders, including but not limited to schizophrenia (18). Deletions and duplications between break points one and two of chromosome 15q11.2, known as the Prader-Willi/Angelman syndrome region have also been found to be associated with an

increase in risk. Additionally, large deletions have been identified that affect risk for schizophrenia on 1q21.1 and 15q13.3 (44), as well as CNVs that increase the risk for schizophrenia, autism, and other psychiatric disorders on 3q29, 16p11.2, and 17q12 (18).

In addition to the effect that CNVs have on risk, another mechanism that has been found to increase risk of sporadic cases of schizophrenia in an Afrikaner population includes the presence of *de novo* mutations. 17.6% of sporadic cases were found to have *de novo* mutations, and 9.9% were found to have *de novo* CNVs (42). Exome sequencing was performed on cases and controls who did not have CNVs, in conjunction with parental studies. It was found that *de novo* variants were 3.7 times more likely to harbor nonsynonymous changes that increased the risk for schizophrenia compared to rare inherited changes (42). Four genes were identified as having higher rates of recurrent *de novo* mutations among cases compared to controls. These genes were *DPYD, TRRAP*, *VPS39*, and *LAMA2*. In addition, five genes were identified to be associated with risk for schizophrenia when altered by both *de novo* mutations and CNVs. These genes include *DGCR2, TOP3B, CIT*, *STAG1*, and *SMAP2* (42).

2.4 HOMOZYGOSITY-BY-DESCENT

Homozygosity of alleles is more likely to occur when some form of inbreeding occurs. If we take the inbreeding level in the population as f, following traditional population genetics practice (30), and take the frequency of the rare allele as *p*, then the frequency of the homozygote, which under Hardy Weinberg equilibrium is p^2, is increased to $p^2 + p(1-p)f$. The impact, in terms of the approximate rate, is $1 + f/p$. If *p* is small relative to *f*, homozygotes of the rare allele occur much more frequently then by chance and these homozygotes will tend to be

homozygous-by-descent (HBD). HBD is different from homozygosity per se because the variants, and typically the regions surrounding them, not only possess similar allele variants, but are actually identical alleles that came from the same founder. Recent literature has focused on individuals who inherit chromosomal segments with long stretches of consecutive homozygous SNPs, called runs of homozygosity (ROHs). Although all HBD subjects have regions with ROH, not all ROH are HBD. The rareness of having two chromosomes that share large segments of identical genetic information provides evidence that those two segments came from the same founder. Shorter segments that share identical information, which are rare in the population, also offers support that those segments are homozygous-by-descent.

Studies have suggested that ROH (and, implicitly, HBD) is associated with risk for schizophrenia in certain samples (10, 11). However, the mechanism generating the association is not obvious and specific gene regions harboring highly penetrant recessive alleles have yet to be identified. According to a study done by Keller et al, a 0.01 increase in the amount of ROH segments within an individual's genome can increase their risk for schizophrenia by 17% (11). When analyzing cases that had been diagnosed with schizophrenia versus unaffected controls, it was found that ROHs were more common among some of the affected cases (12). In 2011, a study was done to specifically analyze the association between the presence of homozygosity-by-descent and the prevalence of schizophrenia in Palau. The study looked for statistically significant differences between the occurrence of HBD regions, the number of HBD regions, and the length of HBD regions among cases and controls. The results showed differences in the length of HBD regions, suggesting that cases possessed significantly longer HBD regions compared to controls (10).

No studies have determined whether or not HBD segments located at specific loci of any given chromosome increase the risk for schizophrenia. To investigate this question, we evaluated whether HBD regions show a pattern consistent with highly penetrant recessive risk alleles that are contained therein. If such regions could be found, they would provide a key step for determining risk alleles involved with schizophrenia.

2.5 RELEVANCE TO GENETIC COUNSELING AND PUBLIC HEALTH

Similar to other complex diseases, schizophrenia places a significant burden on health-care resources. The estimated cost involved in the treatment and maintenance of schizophrenia is approximately $32.5 billion annually (20). This includes costs associated with morbidity, social disability, and associated substance abuse (29). By understanding the genetics of schizophrenia, it may be possible to understand its pathophysiology, leading to the development of more effective and cost reducing treatments (19). In addition, earlier detection of schizophrenia and treatment may alter and improve prognosis (19).

Through the identification of susceptibility genes for schizophrenia, it would be possible to offer genetic testing to at risk individual. Identification of a susceptibility allele would lead to the initiation of preventative medicine, help identify early signs of disease onset, and would aid in determining which family members are also at risk. Currently, the inheritance of schizophrenia is considered to be complex and multifactorial (19). However, it is possible that other modes of transmissions can exist depending on the identified molecular etiology. Given what has been observed in Palau as well as in other isolated populations, it is possible that certain recessive alleles could be associated with risk for schizophrenia. Understanding the inheritance

of schizophrenia can allow for a better assessment of at risk individuals, and provide a better estimate of recurrence risk (21). Individuals who knew they were at risk could then take the steps to consult with a psychiatrist to discuss signs of the disease and expectations regarding prognosis (19).

Knowing that schizophrenia was a genetic condition not only put individuals in more control over their health care management, but it was also found to decrease stigma (21). Individuals may feel that a genetic attribution to schizophrenia moves the blame and responsibility of the disease away from the individual and onto the biological causes. Those who believed that schizophrenia was related to genetics were more likely to comply with pharmacological interventions and recommendations for hospitalizations (21). Part of this was because of the perception that a genetic illness is more serious and permanent, and thus required more attention. In a genetic counseling session, it is important to emphasize both genetic and environmental factors associated with the disease, as well as the psychosocial impacts that the disease can have on patients and their families.

3.0 MATERIALS AND METHODS

3.1 SAMPLE AND GENOTYPING

Continuing with the ongoing study of schizophrenia in Palau, DNA from blood samples of a total of 573 subjects were genotyped using Affymetrix Genomewide Human SNP Array 5.0, which include all known affected subjects (n=208, 36.3%), a sample of relatives of an affected (n=191, 33.3%), and control subjects (n=174, 30.4%). Research protocols and procedures were approved by institutional review boards (IRBs) at each of the sites in the US and the Republic of Palau. All subjects provided written informed consent to participate after receiving a full explanation of the study in both English and Palauan.

Affected individuals were identified by medical records or referral by a family member, and confirmed using a modified version of the Schedule for Affective Disorders and Schizophrenia-Lifetime Version (SADS-L). Interviews were followed by intensive review of psychiatric medical records that described symptoms and treatment of the illness. The SADS-L interviews in combination with medical records helped determine whether a DSM-IV-based diagnosis of schizophrenia or a related psychotic disorder could be achieved. Subjects who did not meet criteria for schizophrenia or a related psychotic disorder (e.g., bipolar disorder, psychosis, schizoaffective) were excluded.

Controls were identified as subjects who had no history of psychiatric illness. In addition, controls had to be distantly related to the cases by at least three degrees. (Note that all individuals in Palau are likely to be related to some degree due to its isolated nature.)

Quality control (QC) was conducted at the individual and SNP levels. At the individual level, samples were excluded if called genotypes fell below 96.2%; discrepancy between nominal and genetically-determined sex; or a large number of Mendelian errors (>20,000). SNPs were excluded if they were unmapped or monomorphic, had a call rate below 98.3% or more than 10 Mendelian errors. After this step, all SNPs passed a test for Hardy-Weinberg Equilibrium and there were 193 cases and 143 controls genotyped for 354,779 SNPs for analysis.

3.2 IDENTIFYING ROH AND HBD SEGMENTS

The analysis toolset PLINK was used to evaluate the presence of ROH, which required a minimum of 100 consecutive homozygous SNPs in order for a stretch to be classified as an ROH (28). The density specified is a minimum of 1 SNP per 50 kb, and there could not be a gap between SNPs that is greater than 2 Mb. To account for error, up to 2 missing genotypes were permissible (i.e., of 102 SNPs in a region, if 100 were homozygous and two interspersed SNPs had missing genotypes, this would be an ROH). We used Beagle to detect regions that are HBD (14). HBD regions were determined based on the rarity of the SNPs that compose a given ROH, and the likelihood that an individual could be homozygous for that particular haplotype had it not come from the same founder (14). A region is classified as HBD if, for each of the consecutive SNPs found therein, the probability that the SNP allele originated from a common ancestor exceeds 0.5. (14). The specific SNP range that the HBD region falls between is measured starting

with the first SNP in which an individual is determined to be HBD and continues along the haplotype until a SNP is reached that has an HBD probability of less than 0.5 (14). The subjects who had stretches that met these requirements were considered subjects with HBD regions.

3.3 SEARCHING FOR PENETRANT RECESSIVE RISK ALLELES

Previous analyses in the Palauan sample found individuals with schizophrenia to have greater total length of HBD compared to controls. However, whether or not these HBD regions harbor highly penetrant recessive alleles remains unclear. To investigate this further, we identified haplotypes within HBD regions that had the potential of being associated with risk for schizophrenia. Previous simulations showed that 3 or more subjects HBD at the same SNP would happen less than 5% of the time if HBD regions were randomly distributed across subjects' genomes. Thus, we examined haplotypes in HBD regions where there were 3 or more cases sharing an HBD region. We evaluate overlap of HBD segments by dividing the region into different segments based on HBD sharing and count the number of cases and controls sharing each segment.

Homozygosity for each SNP between cases and controls were compared using the Modified Quasi-likelihood Score test (M_{QLS}) and a recessive mode of inheritance, with major and minor alleles specified as the recessive allele. This testing provides an appropriate test of the null hypothesis or no association between homozygous SNP (segment) and schizophrenia risk, even when the sampled individuals are known to be related (22). Accounting for relatedness is important because in the Palau sample, genotypes will be more correlated among subjects due to shared ancestry. The relatedness of each subject that was homozygous for a particular segment

was determined with the use of genetic-based relationship matrices (Appendix A). Evaluation of haplotype structures was also performed by hand to evaluate patterns that would be difficult to capture by summary statistics such as those from M_{QLS}. Final conclusions of the relationship between the haplotype stretches found in HBD regions and association with schizophrenia risk were a synthesis of calculations such as M_{QLS} and careful inspection of patterns of haplotype sharing in pedigrees.

4.0 RESULTS

4.1 IDENTIFYING HBD AND ROH SEGMENTS

A total of 432 HBD regions were identified across the genome of the Palauan subjects. Approximately 54.8% of the subjects in Palau carried at least one HBD region, and 28% carried more than one HBD region. The average number of HBD regions per subject is 1.3 (±2.1) and the average length of an HBD region is 5.1Mb (±7.03). The smallest HBD region was 143.9 kb, and the largest HBD region was 78.4Mb. Subjects who had HBD regions had an average HBD length of 11.9 Mb (±19.9). Almost all subjects in the sample carried at least one ROH. The average number of ROH per subject was 72.6 (±20.56). The average length of an ROH in the subjects was 1.34Mb (±14.7Mb). The average cumulative length of ROH covered 3.5% of the genome at 100.5Mb (±40.81).

While evaluating overlap of HBD segments among cases and controls, a significant trend favoring greater case representation was found relative to control representation in loci with higher numbers of HBD segments (Table 2). Cases and controls were also compared for relative homozygosity of each SNP using M_{QLS} with recessive mode inheritance (Figures 1 and 2). In Figures 1 and 2, the top shows comparison of cases and controls on each SNP across the genome with regards to recessive inheritance of the major allele, and the bottom shows the comparison of cases and controls on SNPs in HBD regions shared by three or more cases. None of the SNPs

reached genome-wide significance. However, several regions were identified in which there was overlap among three or more affected subjects (Table 3). Although some subjects had HBD along the same region, they were not necessarily identical-by-descent to each other throughout the entire region. This meant that individually, they were all homozygous in the same region, but they were not all homozygous for the exact same haplotype variant.

Table 2. Number of shared HBD regions in cases and controls

Number of controls	Number of cases				
	4	3	2	1	0
0	2	19	82	202	-
1	1	12	46	106	130
2	0	2	7	15	35
3	0	1	0	2	6
4	0	0	0	0	0

Ordering the counts for cases:controls shared regions by 1:0 *vs.* 0:1, 2:1 *vs.* 1:2, 2:0 *vs.* 0:2, 3:2 *vs.* 2:3, 3:1 *vs.* 1:3, 3:0 *vs.* 0:3; 4:1 *vs.* 1:4; 4:0 *vs.* 0:4), there was an excess of cases sharing HBD regions as compared to controls (Chi-square=8.97, df=1, p=0.0027).

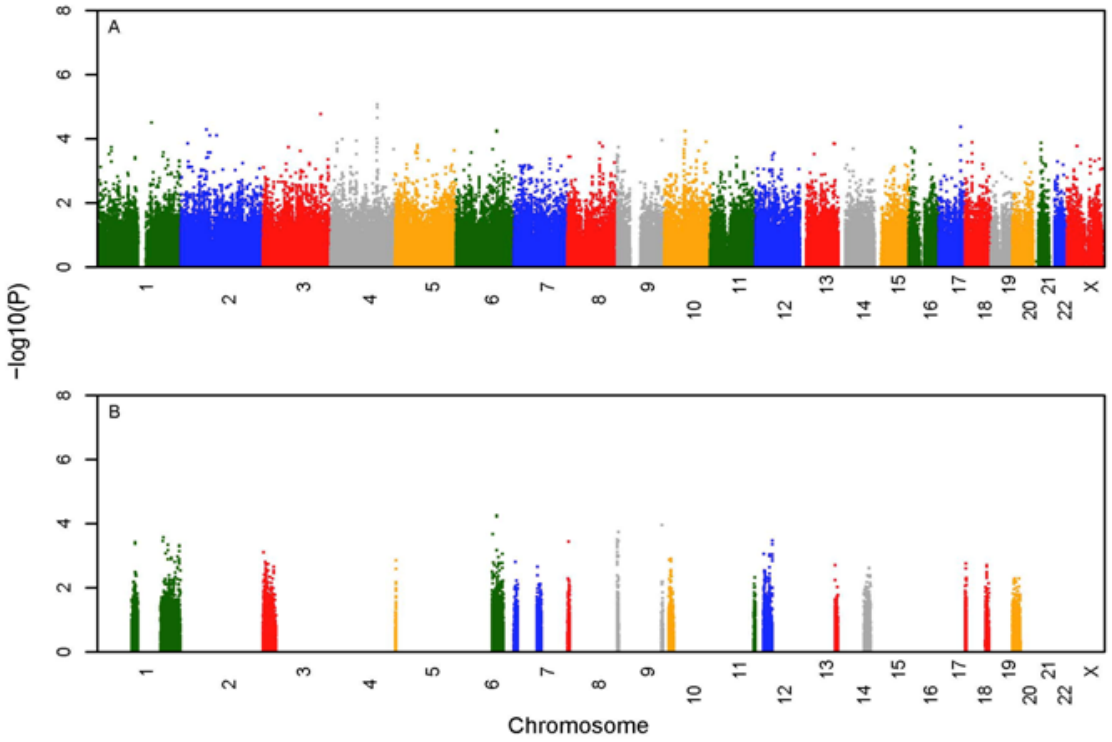

Figure 1. Comparison of cases vs. controls on homozygosity for SNPs using MQLS with recessive mode inheritance (major allele)

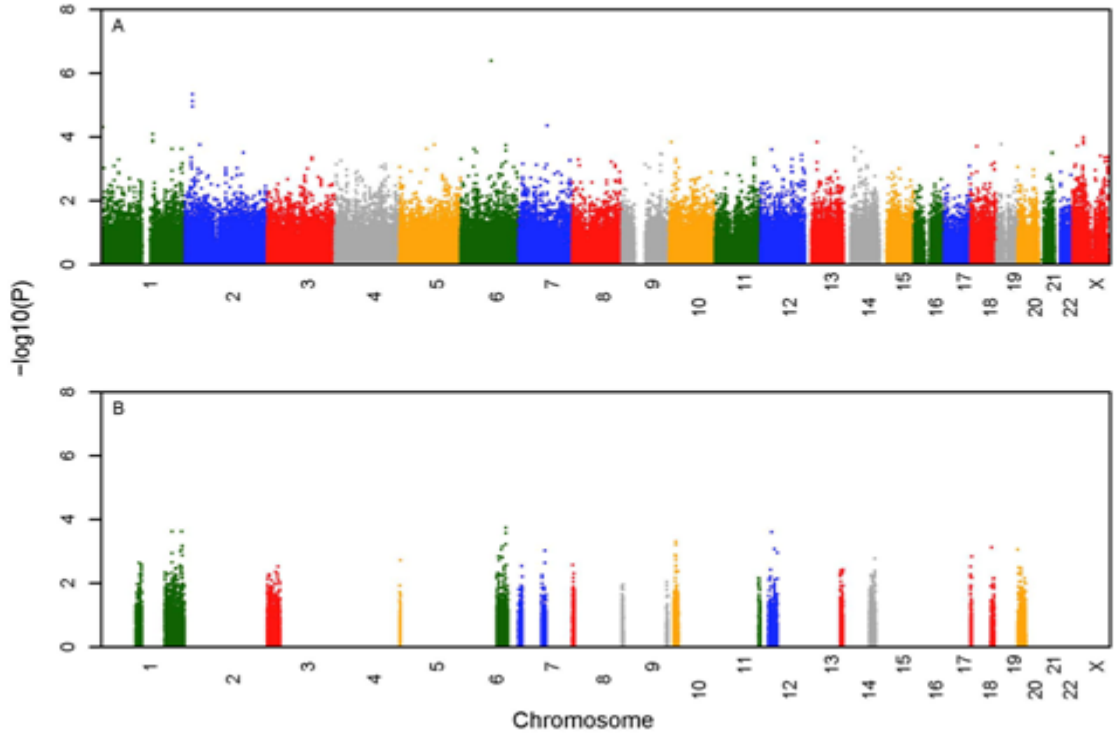

Figure 2. Comparison of cases vs. controls on homozygosity for SNPs using MQLS with recessive mode inheritance (minor allele)

Table 3. HBD regions with 3 or more affected subjects

Chromosome	SNP Region	Number of HBD subjects
1	107519211-111117468	3
1	206897233-222097120	4
1	227316809-231762016	3
3	8484476-13288817	3
5	1514166-2622710	3
6	106922247-124024764	3
6	128753042-133915178	3
7	792019-6151881	3
7	9144986-10546273	3
7	72951791-79981363	4
8	3165317-6414803	2*
9	3236490-5398357	4
9	134128820-135954270	3
10	15111853-21607729	3
11	130376165-134443679	3
12	25105886-33738526	4
13	105576622-109061569	3
14	80741858-83254290	4
18	208084-2062155	3
18	63511136-69498612	3
20	5199068-12227537	4

*although the HBD region on chromosome 8 only had two affected subjects that overlapped that region, it was included in the analysis because both subjects were identical-by-descent through the entire region. This did not end up being a significant finding.

4.2 SEARCHING FOR PENETRANT RECESSIVE RISK ALLELES

Our strategy for identifying highly penetrant recessive risk alleles included examining the genotypes of affected sibling pairs, as well as cases and controls. Of the affected sibling pairs in our sample, there were three regions of the genome shared HBD, and one sibling pair had two of the HBD regions. We then evaluated other subjects, both affected and unaffected, to identify shared homozygous segments of these regions. At a minimum, we considered sub-haplotypes of the original HBD haplotype of at least length 10 SNPs. Upon evaluating these regions for additional subjects who share such haplotype (Table 4) one haplotype consisting of 34 SNPs

(chromosome 2: 187,577,486-187,986,279) was shared by an additional 18 cases and 5 controls, giving an odds ratio of 2.87.

Although the odds ratio does not lend much support to the possibility of a highly penetrant recessive risk allele in the region, we investigated this issue further by looking at extended haplotypes of these 18 cases and 5 controls (chr2: 186,973,280-188,261,318) (Figure 3). Several haplotypes were identified and labeled A, B, D, and E. We labeled the shared haplotype between the siblings sharing the region HBD, as haplotype H. Two additional pairs of siblings in separate nuclear families, were determined to be homozygous for haplotypes A and B, and were assumed to be identical-by-descent. Additional identified haplotype variants had slight differences compared to these haplotypes, which we label as A', A", H', D, E, and E'.

Table 4. Homozygous haplotypes shared with siblings with HBD sharing

ID	Affection	Haplotype Chr2: 186,973,280-188,261,318	Haplotype Index
Siblings sharing HBD region			
20306*	Affected		H
			H
22354*	Affected		H
			H
Cases and controls sharing haplotype with HBD siblings			
19485	Affected		A
			A
20310*	Affected		A
			A
20330*	Affected		A
			A
22372	Affected		A
			A'
22396	Affected		A
			A"
22476	Affected		A
			H'
20302*	Affected		B
			B
25057*	Affected		B
			B
20293*	Affected		A
			B
22361	Affected		H
			B
20291	Affected		A
			D

Table 4. Continued

22369	Affected	2112111112121122222211111211221211121212222212212222111112112111212112122121221222112222112122212 12212	B
		211111111112121111111122221121212212121222221221222211111211211121211212212122122211222211212221212212	A
22358	Affected	12122212111121222222212112222212212121222221221222211111211211121211212212122122211222211212221212212	E
		211111111112121111111122221121212212121222221221222211111211211121211212212122122211222211212221212212	A
20271	Affected	12122212111121222222212112222212212121211121221222211111211211121211212212122122211222211212221212212	E'
		21121111121211222222112122121212212121222221221222211111211211121211212212122122211222211212221212212	D
22530	Affected	2111111111121211111111222211111212212121111222212112111112112111212112122121221222112222112122212 12212	
		2112111112121122222211111211221211121212222212212222111112112111212112122121221222112222112122212 12212	B
14729	Affected	12122212111121222222212112222212212121222221221222211111211211121211212212122122211222211212221212212	E
		11111111111112111111121221121212212121222221221222211111211211121211212212122122211222211212221212212	
3915	Affected	21121111121111111111122221121212212121222221221222211111211211121211212212122122211222211212112112212	
		12122212111121222222212112222212212121222221221222211111211211121211212212122122211222211212221212212	E
17624	Affected	211211121212112222222112122112212212121222221221222211111211211121211212212122122211222211212221212212	
		1212111111111212222222111112111221211212122222122122221111121121112121121221212212221112 22111121112221122	
26075	Controls	1222112121211122222221111121122121112121222221221212211111211211121211212212122122211222211212221212212	H
		2112111112121122222211111211221211121212222212212222111112112111212112122121221222112222112122212 12212	B
26105	Controls	211111111112121111111122221121212212121222221221222211111211211121211212212122122211222211212221212212	A
		2112111112121122222211111211221211121212222212212222111112112111212112122121221222112222112122212 12212	B
26172	Controls	21121111121211222222112122121212212121222221221222211111211211121211212212122122211222211212221212212	D
		211111111112121111111122221121212212121222221221222211111211211121211212212122122211222211212221212212	A
60103	Controls	2112111112121122222211111211221211121212222212212222111112112111212112122121221222112222112122212 12212	B
		2111111111111211111112222211121212212121222221221222211111211211121211212212122122211222211212112112212	
60012	Controls	2112111112121122222211111211221211121212222212212222111112112111212112122121221222112222112122212 12212	B
		21121111121211222222112122121212212121222221221222211111211211121211212212122122211222211212221212212	D

Subjects with IDs in bold and similar letter superscripts are siblings; Bold SNPs are the 34-SNP haplotype shared by all subjects

We examined haplotypes for family members of affected and control individuals and found the distribution of these haplotypes in families not to be consistent with the segregation of a penetrant recessive risk variant. We present in Figure 3 the pedigrees and haplotype sharing for the three nuclear families with pairs of siblings sharing homozygous haplotypes assumed to be identical-by-descent. In Family 1, the affected siblings (20306, 22354) are homozygous for H. In Family 2, the affected siblings (25057, 20302) are homozygous for B. In family 3, the affected siblings (20330, 20310) are homozygous for A. In Family 2, there are two additional siblings (25071, 25073) that are unaffected, yet homozygous for B similar to their affected siblings. In Family 3, there are two affected siblings (13872, 20256) that are heterozygous for A. Based on haplotype structure, haplotypes A and B show 67% and 86% similarity with haplotype H, which is a portion of the original HBD segment. Thus, there was no consistent pattern that offered support that there was association between the haplotypes and disease status.

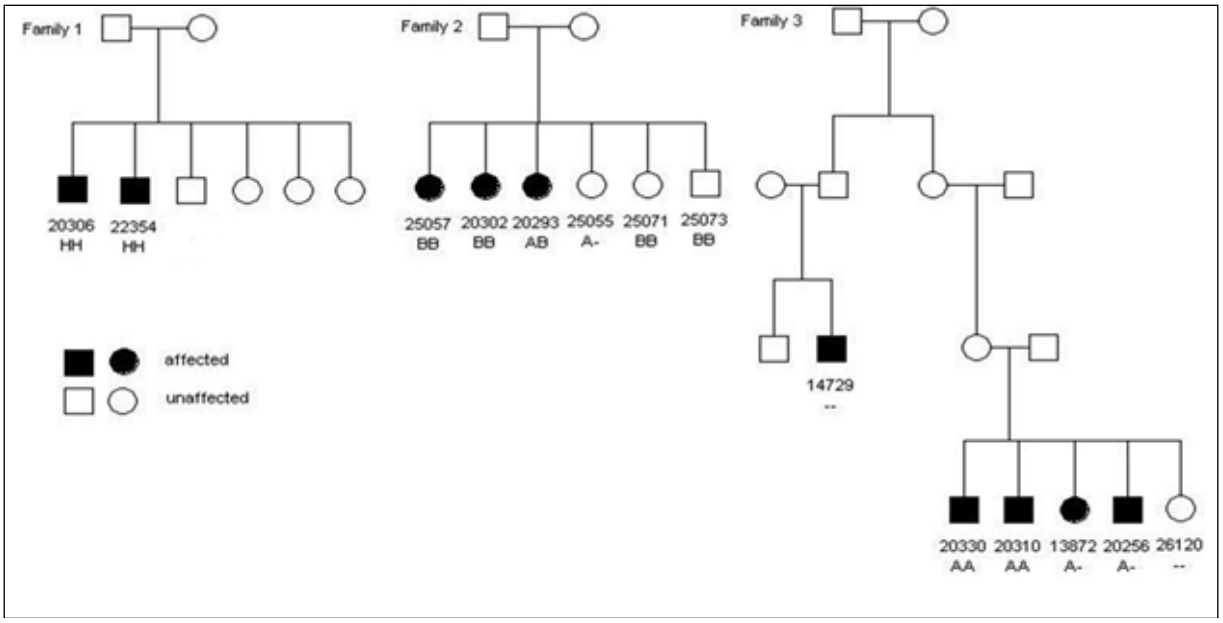

Figure 3. Family pedigrees of affected siblings with shared HBD regions
Symbols without subject ID numbers did not have genotyping information

In addition to examining HBD regions shared by siblings, sub-haplotypes within the larger HBD haplotypes shared by three or more cases were evaluated. Evaluation of sub-haplotypes amongst additional cases and controls revealed some sub-haploptypes that were overrepresented in cases, relative to controls. Five particular sub-haplotypes were evaluated to identify if any were potentially harboring recessive risk alleles. These regions overlapped with genes *RAB3GAP2*, *MARK1*, *DISC1*, *FAM184A*, and *IGSF9B. RAB3GAP2* had previously been implicated in autosomal recessive diseases (24), and *MARK1* and *DISC1* had previously been associated in risk for autism and schizophrenia (23, 25). This information is summarized in Tables 5 and 6.

Table 5. Summary of Cases and Controls who are homozygous for the haplotype

Chromosome	HBD region	SNP stretch	Gene	Cases Homozygous for Haplotype	Controls Homozygous for Haplotype	Odds Ratio
1	206897233-222097120	218368739-218511547	RAB3GAP2*	14	8	1.3296
1	206897233-222097120	218579529-218906324	MARK1**	9	3	2.2989
1	227316809-230388213	230223388-230384737	DISC1***	7	2	2.6720
6	106922247-124024764	118587907-119427882	FAM184A	7	1	5.3817
11	130376165-134443679	133954456-134090010	IGSF9B	14	7	1.5307

*Previously implicated in autosomal recessive disease
**Previously implicated in autism
***Previously implicated in schizophrenia

Table 6. Summary of Cases and Controls with one copy of the haplotype

Chromosome	HBD region	SNP stretch	Gene	Cases with one copy of the Haplotype	Controls with one copy of the Haplotype	Odds Ratio
1	206897233-222097120	218368739-218511547	*RAB3GAP2*	91	70	0.9431
1	206897233-222097120	218579529-218906324	*MARK1*	62	39	1.2742
1	227316809-230388213	230223388-230384737	*DISC1*	53	37	1.0948
6	106922247-124024764	118587907-119427882	*FAM184A*	50	36	1.0490
11	130376165-134443679	133954456-134090010	*IGSF9B*	84	52	0.9391

4.2.1 RAB3GAP2 Region

Four affected subjects shared a locus HBD on 1q41 comprised of 1,695 SNPs spanning 206,897,233 bp to 222,097,120 bp (Build HG 18). None of these subjects, however, share a substantial portion of their haplotypes. The largest portion of the locus shared among three of the

subjects is an 11 SNP sub-haplotype (218,368,739 bp to 218,511,547 bp). (Figure 4-A) This sub-haplotype falls within the RAB3GAP2 gene. For this sub-haplotype, we identified 14 cases and 8 controls that were also homozygous and yielding an odds ratio of 1.33.

The odds ratio does not lend much support for a highly penetrant recessive risk allele in this region. Thus, it is possible that sharing of this 11 SNP sub-haplotype could have been due to a risk allele with a different inheritance pattern, such as an additive effect on risk. To evaluate this possibility, we looked through all variants of the haplotype located at this 11 SNP locus that were carried by cases. Two variants of the haplotype (Figure 5) were found in homozygous or compound heterozygous cases. Cases were either homozygous for one of the variants, or compound heterozygous for both. Interestingly, these two variants were composed similarly, differing only by two SNPs. We then expanded the haplotype 20 SNPs to the right and 30 SNPs to the left to examine a larger 61 SNP haplotype that ranged from SNP location 218,177,182 to 218,725,045. There were 206 variants of this extended haplotype found among subjects. Each variant was assigned a number 1-206. Of the 206 variants, fifteen were found among five of the nuclear families in the study (Figure 6).

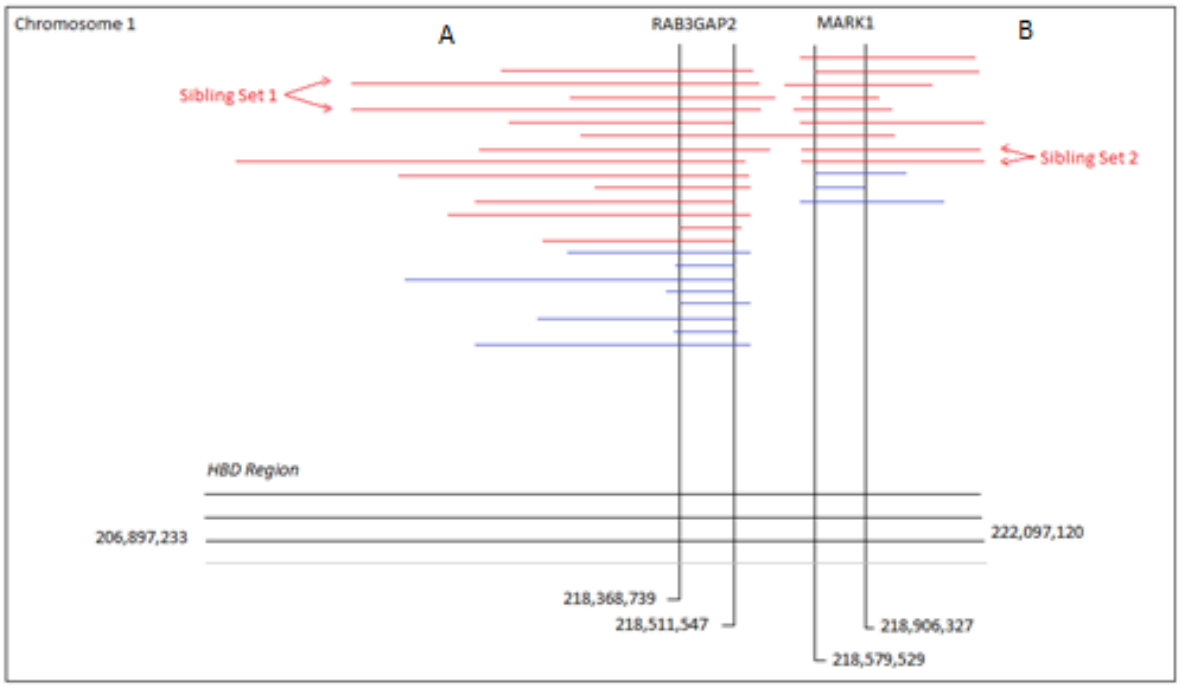

Figure 4. Association Analysis of RAB3GAP2 (A) and MARK1 (B) regions

SNPs of Haplotypes (218368739-218511547)										
1	1	2	1	2	2	1	1	1	2	1
1	2	1	1	2	2	1	1	1	2	1

Figure 5. The original RAB3GAP2 haplotype stretch

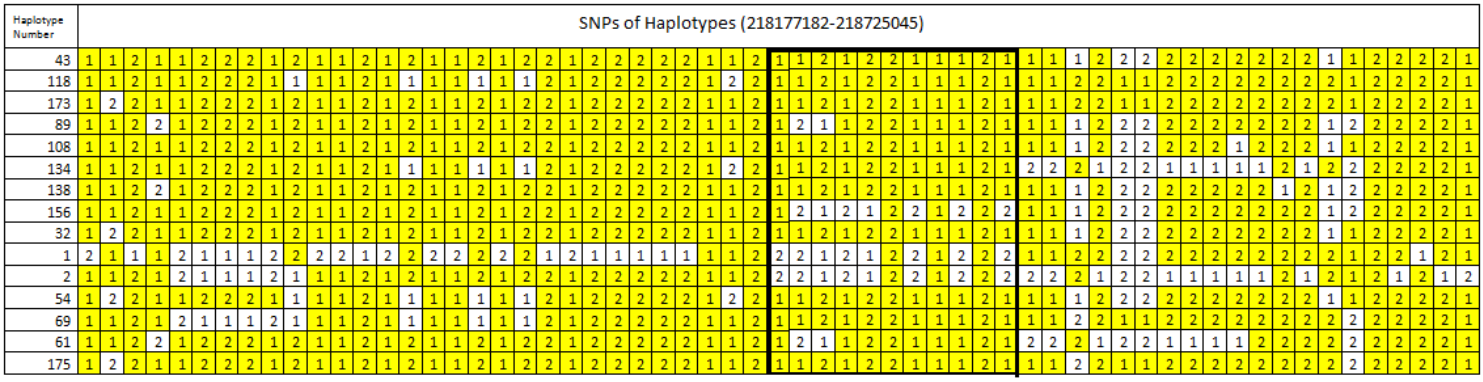

Figure 6. The extended RAB3GAP2 haplotype stretch

Pedigrees of the five families (Figure 7) display who in each family carries which haplotype variant. Black symbols represent family members who are affected by schizophrenia. None of the haplotype variants appeared to be consistent with highly penetrant risk alleles for schizophrenia. Variant 43 was initially considered as a potential risk variant because it was identified in three cases from Family 1 (13881, 20256, 20310), and no controls. This variant is very rare in the population, having a frequency of less that 0.03% and only appearing in the three individuals from Family 1. However, because of its low frequency and appearance in only this one family, it is likely that it is a variant unique to the family and not associated with schizophrenia in the Palau population. More evidence would have to be provided to suggest that it is related to the disease status of the subjects.

Variant 134 appears in two subjects from Family 2 as a compound heterozygote with variant 118. However, the mother of the family is unaffected and the son is affected despite having identical genotypes in that region. This inconsistency allows us to conclude that those variants are not associated with disease status in the family.

Variant 32 is one of the more common haplotypes and appears in approximately 2% of the Palauan population. It also appears in two of our families: Family 3 and Family 5. In Family 3, subject 22420 is affected with schizophrenia and is homozygous for variant 32. His brother,

subject 26154 is only heterozygous for the variant. Initially, this hinted that variant 32 may be related to schizophrenia when in the homozygous state, following a recessive inheritance pattern. However, when variants 32 and 118 are compared based on the SNPs that compose the haplotypes, there is very little difference. Only ten of the 61 SNPs are different and there are still long stretches within the haplotype where there is consistent SNP sharing. For this reason, there is not enough evidence to conclude that homozygosity of variant 32 is the reason for schizophrenia expression. In Family 5, two siblings (26165 and 5000) share the same 175/32 genotype, yet only subject 5000 has positive disease status. This suggests that this particular genotype is unrelated to having schizophrenia.

Variant 54 is also found in 2% of the population, but in only one of the families. The affected subject, 22522 is a compound heterozygote for variants 54 and 69. When looking at these specific haplotypes, there are several breaks in SNP sharing. Because of this, it is unlikely that the subject would be affected by schizophrenia because of homozygosity in this region. Without further evidence, it cannot be said that variant 54 a has significant relation to disease status.

Variant 175 is a rare variant with a frequency of 0.08% in the Palauan population. It is found in three individuals in Family 5, two of which are affected. However, just as with variant 32, the two brothers in the family have the same genotype but opposite disease status. This suggests that the components of the genotype are not contributing to their schizophrenia status.

In support of our findings, we coded the 11-SNP and 61-SNP haplotypes into an allele format (1 has haplotype of interest, 2 does not have the haplotype) to run M_{QLS} and examine haplotype frequencies in cases and controls taking into account relatedness. No significant

differences were found between cases and controls for the 11-SNP (p=0.394) and 61-SNP haplotype frequencies (p=0.434).

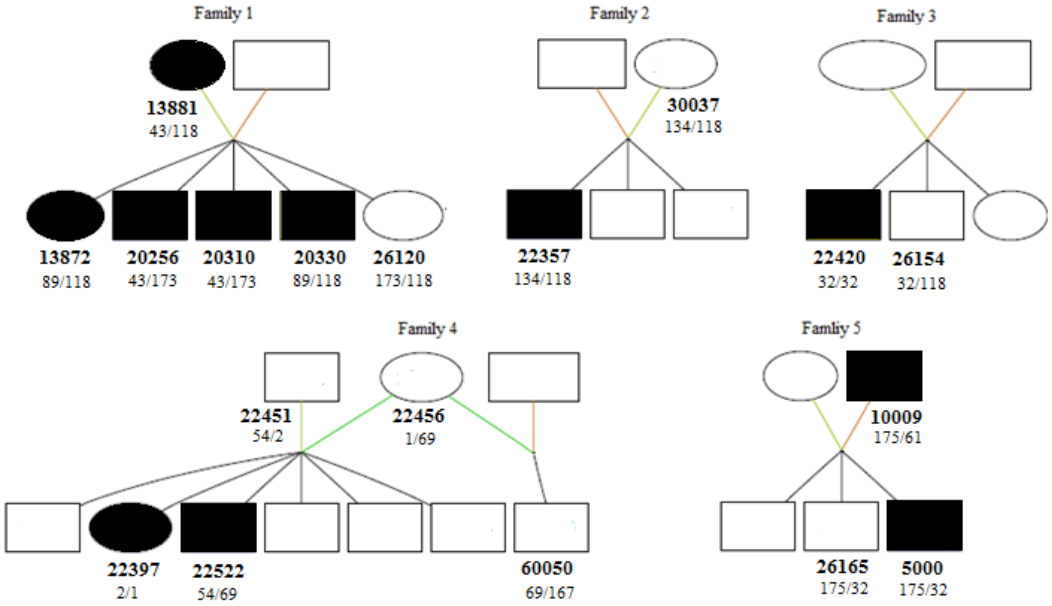

Figure 7. RAB3GAP2 family pedigrees

4.2.2 MARK1 Region

Adjacent to the *RAB3GAP2* region that is within the HBD locus on 1q41, there is another portion of the locus that is shared among three of the subjects. This includes a 16 SNP sub-haplotype (218,579,529 bp to 218,906,324 bp). (Figure 4-B). This sub-haplotype falls within the *MARK1* gene. For this haplotype, we identified 9 cases and 3 controls that were also homozygous and yielding an odds ratio of 2.30.

Although this odds ratio suggested that over twice as many cases would be homozygous for this sub-haplotype compared to controls, we wanted to evaluate further whether this locus was harboring highly penetrant recessive risk alleles. To investigate this, we looked through all variants of the haplotype located at this 16 SNP locus that were carried by cases. There were 51

variants of the haplotype identified. Each variant was assigned a number 1-51. Of the 51 variants, five were found among four of the nuclear families in the study (Figure 8.)

Haplotype number	SNPs of Haplotypes															
27	2	1	1	2	2	2	2	1	1	1	2	2	2	1	1	1
6	2	1	2	1	2	1	2	1	1	1	2	2	2	1	1	1
26	2	2	1	2	2	2	2	1	1	1	2	2	2	1	1	1
47	2	1	2	2	2	2	2	1	1	1	2	2	2	1	1	1
4	2	2	2	1	2	1	2	1	2	2	1	1	1	1	1	1

Figure 8. MARK1 haplotype stretch

Pedigrees of the four families (Figure 9) display who in each family carries which haplotype variant. Variant 27 was suspected to follow a pattern that suggested recessive inheritance in that is appears in all four of the families, and every individual who is homozygous for this specific haplotype is also affected with schizophrenia. However, it has a population frequency of close to 17%, so it is not surprising that it is present in all the families. Subjects who carry one copy of variant 27 are unaffected. In both Family 3 and Family 4, it appears that both parents are unaffected carrier of haplotype 27. Their children then inherited both copies of the haplotype and are affected, With this information alone, it was hypothesized that haplotype 27 was a potential disease causal variant that followed the rules of a recessive inheritance pattern.

In order to evaluate this haplotype more closely, the specific SNPs that made up these five haplotypes were analyzed and compared (Figure 8). If the hypothesis were to be correct, then affected subjects should possess haplotypes with significantly more SNP sharing between their two haplotypes than those individuals who are unaffected. Given the composure of the haplotypes, this does not seem to be the case. The heterozygous parents in Family 3 and the heterozygous father in Family 4 all have the 27/47 genotype and are all unaffected. However,

when looking at the difference in the haplotypes between variants 27 and 47, it can be seen that they only differ by one SNP. For their genotype to be causative of the disease status, this one SNP would have to be the reason, which is highly unlikely. The difference in their haplotypes and that of their children is not enough that it should affect whether or not they have schizophrenia. In addition, subject 22587 in Family 1 is also affected, yet the SNPs that make up variants 26 and 6 differ more than variants 27 and 47, and thus would not be considered to have homozygosity in this region.

In support of our findings, we coded the 16-SNP haplotype into an allele format (1 has haplotype of interest, 2 does not have the haplotype) to run M_{QLS} and examine haplotype frequencies in cases and controls taking into account relatedness. No significance difference was found between cases and controls for the 16-SNP haplotype frequencies (p=0.384).

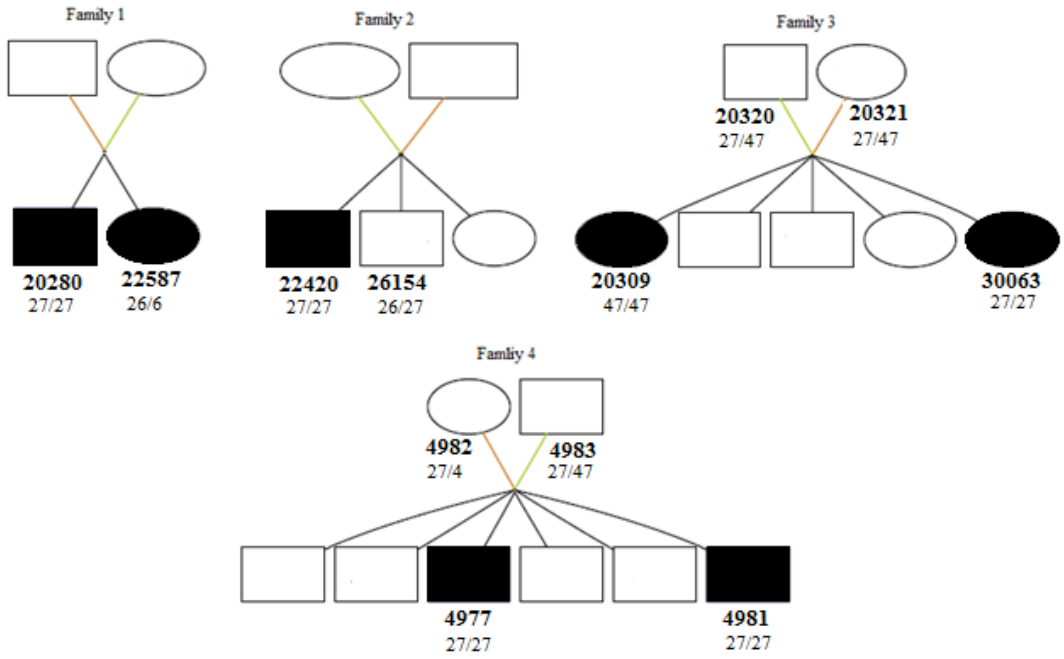

Figure 9. MARK1 family pedigrees

4.2.3 DISC1 REGION

Three affected subjects shared a locus HBD on 1q42.2 comprised of 593 SNPs spanning 227,316,809 bp to 230,388,213 bp (Build HG 18). None of these subjects, however, share a substantial portion of their haplotypes. The largest portion of the locus shared among two of the subjects is a 33 SNP sub-haplotype (230,223,388 bp to 230,384,737 bp.) (Figure 10). This sub-haplotype falls within the *DISC1* gene. For this sub-haplotype, we identified 7 cases and 2 controls who were also homozygous and yielding an odds ratio of 2.67.

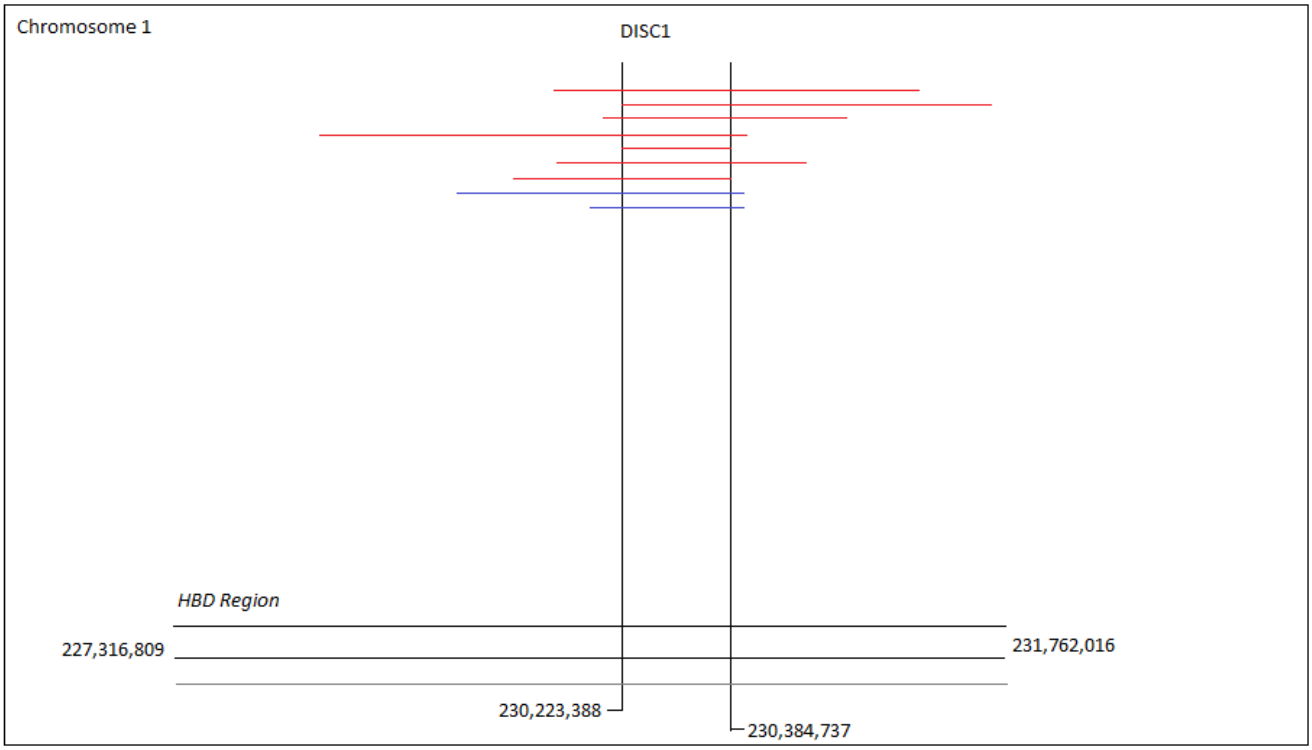

Figure 10. Association Analysis of DISC1 region

Similar to the *MARK1* region, this odds ratio suggests that this haplotype would be homozygous in twice as many cases compared to controls. However, we wanted to evaluate this further before concluding that this locus was harboring recessive risk alleles. To investigate this, we looked through the variants of the haplotype located at this 33 SNP locus carried by cases.

There were 109 different variants of the haplotype identified. Each variant was assigned a number 1-109. Three of the 109 variants were found in one of the study's families. These variants were labeled as 85, 31, and 97 (Figure 11).

Haplotype number	SNPs of Haplotype (230,223,388-230,384,737)																																
85	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	2	1	2	1	2	1	2	1	2	2	2	2
31	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	2	1	2	1	1	2	2	2	2	1	1	2
97	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	2	1	1	1	2	1	2	1	2	1	2	1	2	1	2	2	2	2
97	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	2	1	1	1	2	1	2	1	2	1	2	1	2	1	2	2	2	2
97	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	2	1	1	1	2	1	2	1	2	1	2	1	2	1	2	2	2	2
31	1	2	1	2	1	2	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	2	1	2	1	1	2	2	2	2	1	1	2

Figure 11. DISC 1 haplotype stretch

The pedigree of the family (Figure 12) display who in the family carries which haplotype variant. The limited family history by itself was not able to offer support that these variants were significant to disease risk. However, the composure of these haplotypes was analyzed further, and a pattern was revealed in a smaller 19 SNP stretch (230,260,242 bp to 230 351, 829 bp; highlighted in Figure 11). Variants 85 and 31 were identical in this stretch, making subject 20306 homozygous in this region. Subject 22354 is homozygous as well since he carries two copies of variant 997. Both 20306 and 22354 are affected with schizophrenia. The only unaffected individual is subject 22507 who would be considered heterozygous because 31 and 97 are not similar haplotype variants. This may be important information; however without a larger number of subjects who possess these variants, it cannot be said with confidence this is significant. The 33-SNP haplotype was coded to run M_{QLS}, and no significant differences were able to be found between cases and controls (p=0.785).

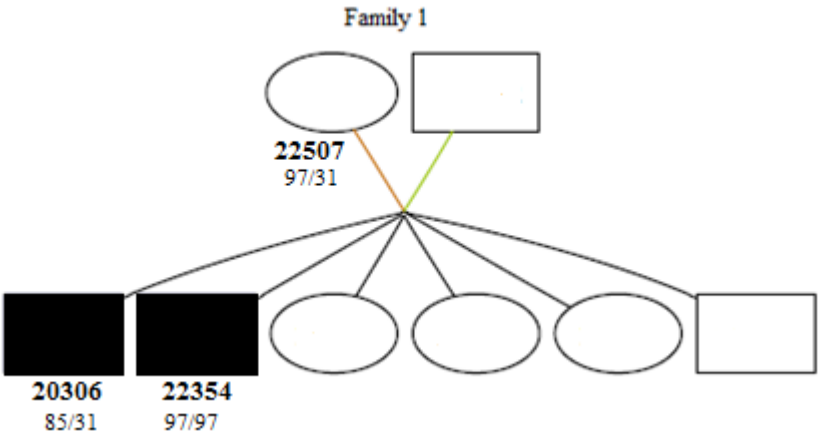

Figure 12. DISC 1 family pedigree

4.2.4 FAM184A REGION

Three affected subjects shared a locus HBD on 6q22.31 comprised of 1,553 SNPs spanning 106,922,247 bp to 124,024,764 bp (Build HG 18). None of these subjects showed sharing in a substantial portion of their haplotypes. The largest portion of the locus shared among two of the subjects is an 82 SNP sub-haplotype (118,587,907 bp to 119,427,882 bp) (Figure 13). This sub-haplotype falls within the *FAM184A* gene. For this sub-haplotype, we identified 7 cases and 1 control that were also homozygous, yielding an odds ratio of 5.38.

This odds ratio is large enough to make this region worth of investigating for highly penetrant recessive risk alleles. Because of this, we looked through all variants of the haplotype located at this 82 SNP locus that were carried by cases, and 151 variants were identified. These variants were labeled with numbers 1-151. Five of the 151 variants were found among two of the study's families (Figure 14). The variants were labeled 2, 136, 65, 146, and 94.

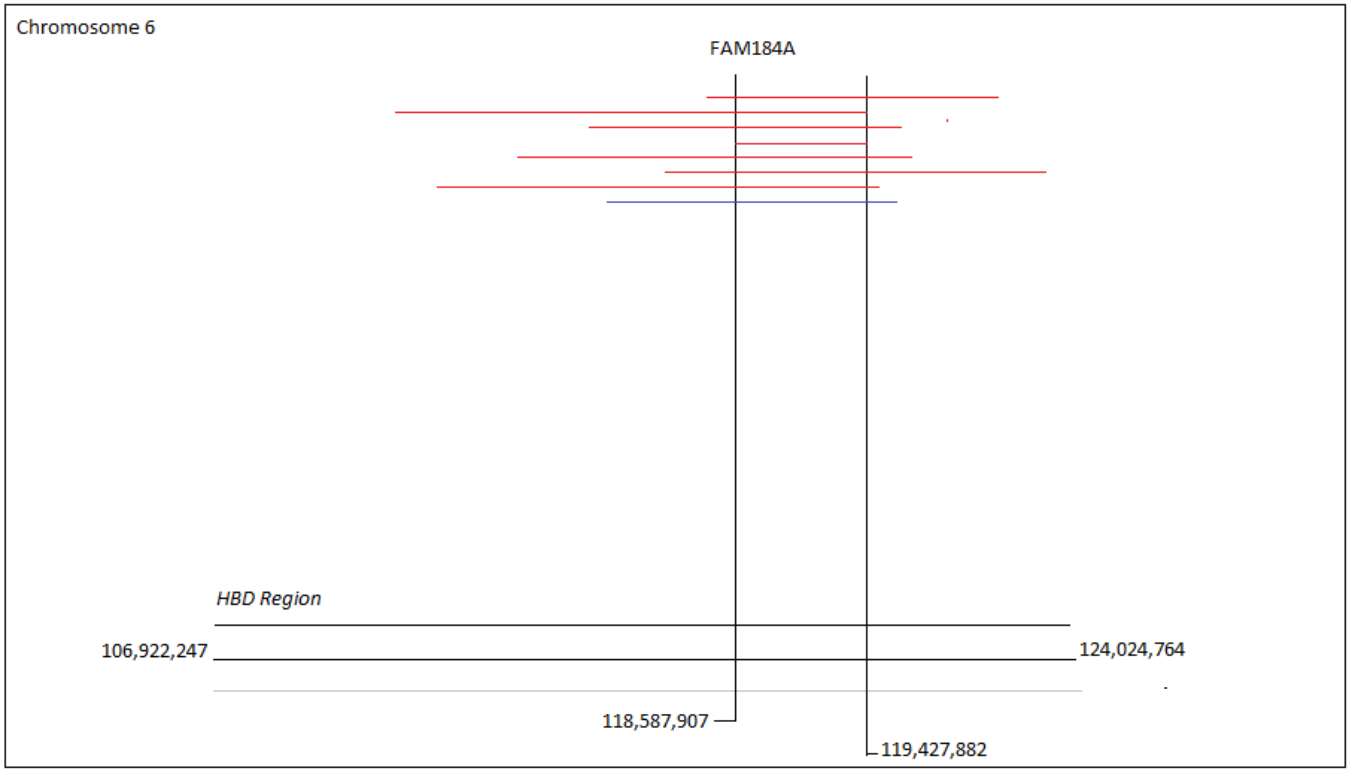

Figure 13. Association Analysis of FAM184A region

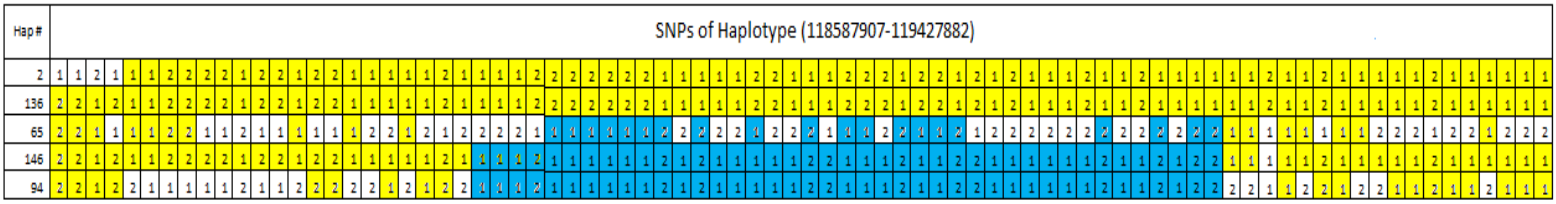

Figure 14. FAM184A haplotype stretch

Pedigrees of the two families (Figure 15) display who in each family carries each haplotype variant. Haplotype variant 136 is homozygous in two of the four affected subjects (22392 and 20306), and heterozygous in subject 22391. If variant 136 was associated with risk, it would be expected that variant 146 would have similar SNP composure to 136 since subject 22391 is also affected and is a compound heterozygote for haplotypes 146 and 136. This is not the case. Subject 22354 is also affected. If homozygosity were important in this region in relation to schizophrenia, it would be hypothesized that 65 and 2 would have similar composure among the

SNPs as well. Again, this is not the case. Coding the 82-SNP haplotype to an allele format to run M_{QLS} supports our findings. No significant differences were found between cases and controls (p=0.216).

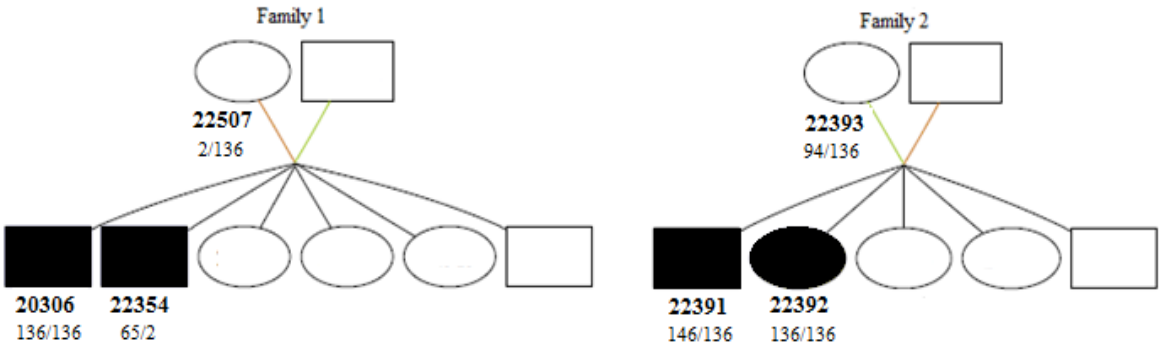

Figure 15. FAM184A family pedigrees

4.2.5 IGSF9B REGION

Three affected subjects shared a locus HBD on 11q25 comprised of 589 SNPs spanning 130,376,165 bp to 134,443,679 bp (Build HG 18). Although none of these subjects shared a substantial portion of their haplotype, they all shared a 24 SNP sub-haplotype (133,954,456 bp to 134,090,010 bp) (Figure 16). This sub-haplotype falls within the *IGSF9B* gene. For this sub-haplotype, we identified 14 cases and 7 controls that were homozygous, yielding an odds ratio of 1.53.

Similar to the *RAB3GAP2* region, this odds ratio does not lend much support for highly penetrant recessive alleles in this region. We evaluated the possibility that sharing of this 24 SNP sub-haplotype could have been due to a risk allele with a different inheritance pattern. As with the other regions, we looked through all variants of the haplotype located at this 24 SNP locus that were carried by cases, and 45 variants were identified. These variants were labeled with

numbers 1-45. Five of the 45 variants were seen among six of the study's families (Figure 17). These variants were labeled 27, 12, 14, 39, and 21.

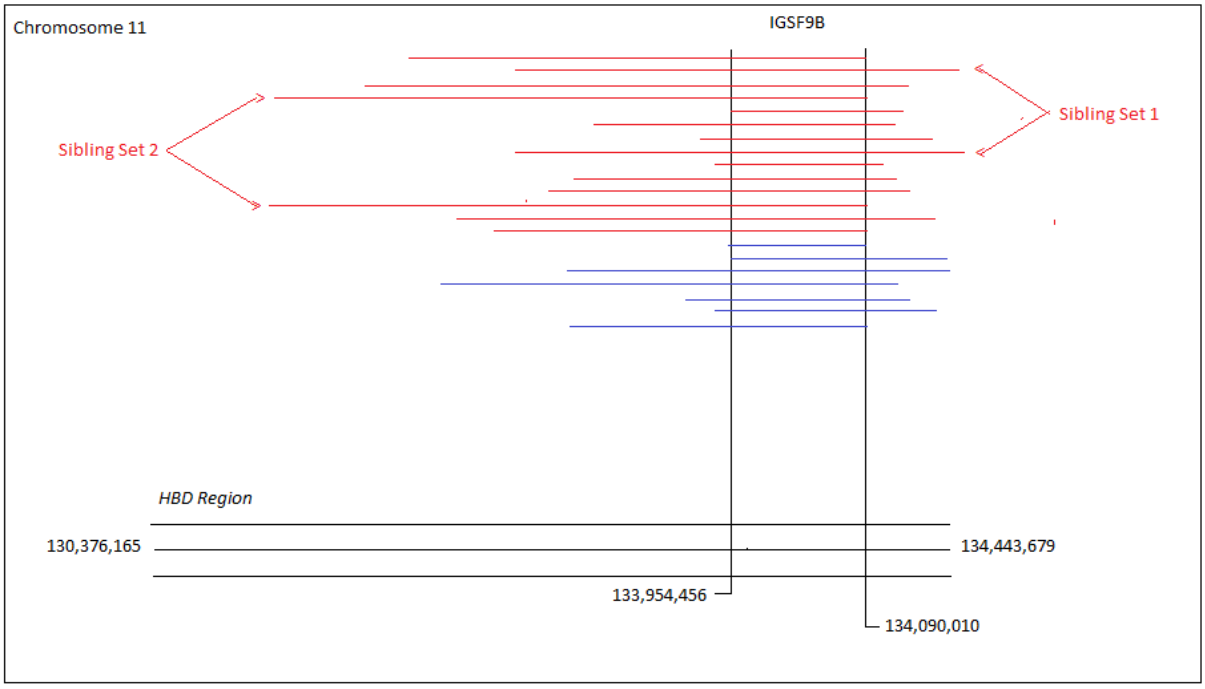

Figure 16. Association Analysis of IGSF9B region

Haplotype Number	SNPs of Haplotype (133954456-134443679)																							
27	2	2	2	2	2	2	2	2	2	1	1	1	2	2	1	2	1	1	1	2	2	1	2	2
12	2	1	1	1	2	1	1	1	1	2	2	2	2	2	1	1	1	2	1	1	1	1	2	2
14	2	1	1	1	2	1	1	1	1	2	2	2	2	2	1	2	1	1	1	2	2	2	1	1
39	1	1	1	1	1	2	2	2	2	1	1	1	1	1	2	1	2	1	2	1	2	2	2	2
21	2	1	1	1	2	1	1	1	1	2	2	2	2	2	1	1	1	2	1	1	1	1	1	1

Figure 17. IGSF9B haplotype stretch

Pedigrees of the six families (Figure 18) display who in each family carries each haplotype variant. Variant 27 is present more often in the families than any of the others. However, it does not appear to show any association with gene status, given that there are an equal number of affected and unaffected individuals who are homozygous for the variant. None of the other variants show a pattern through the families that would be consistent with an association with schizophrenia risk.

To support the findings, we coded the 24-SNP into an allele format to run M_{QLS}. No significant differences were found between cases and controls (p=0.514).

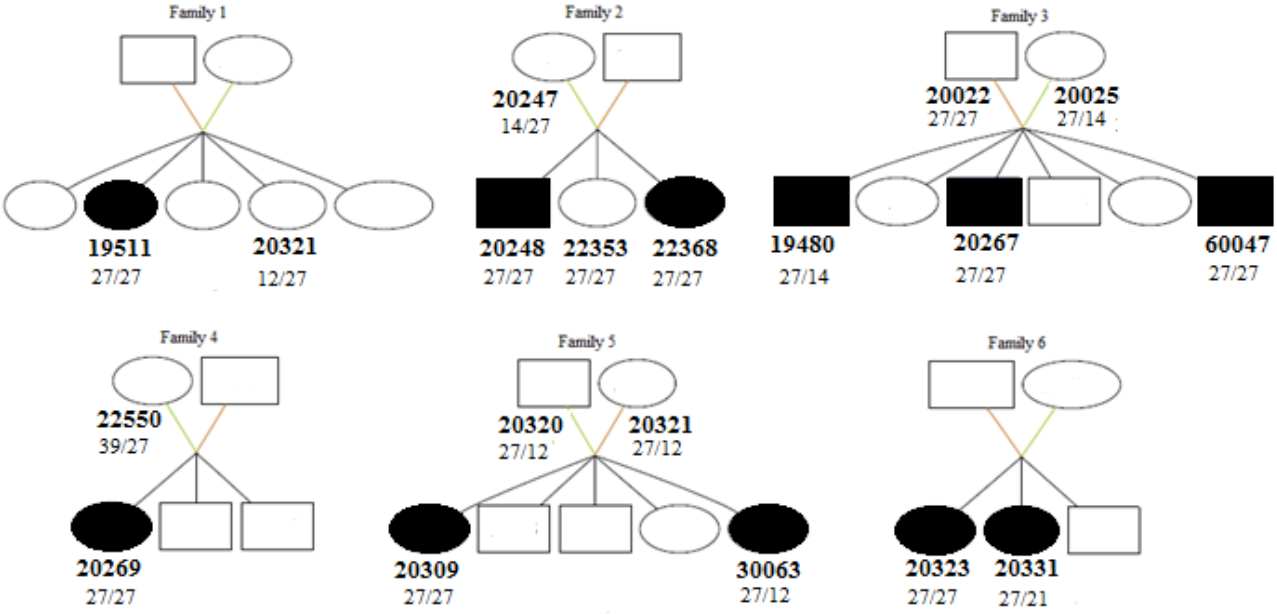

Figure 18. IGSF9B family pedigrees

5.0 DISCUSSION

The heightened prevalence of schizophrenia in the Palau population may be contributed to by its genetic history built from a small founder population, series of bottleneck events, and geographical isolation. Not surprisingly, extensive regions of homozygosity have been found along the genome in the Palauan population. The features of homozygous segments, as described in this study, have been investigated in an attempt to determine whether there is an association between excess regions of homozygosity and risk for schizophrenia. Specifically, our aim was to uncover potential loci that may be harboring highly penetrant recessive risk alleles.

Through an intricate search of homozygous blocks along the genome, we identified regions in which more cases than controls shared HBD segments. These regions overlapped several genes that could be associated with psychosis, due to their expression within functional parts of the brain. *RAB3GAP2*, for instance, has been found to play a key role in neurodevelopment, with expression specifically in the central nervous system (36). Mutations in this gene has also been identified as being the causative of the recessive condition Martsolf syndrome (24). Overexpression of the *MARK1* gene has been identified in the prefrontal region of postmorten brain tissue in individuals with autism. In addition, overexpression of this gene has found to be in association with shortened denderites in animal models (37). *DISC1* has previously been suggested to have allelic variants associated with schizophrenia in the Finnish population, and is believed to regulate multiple aspects of embryonic and adult neurogenesis

(25). Although very little is known about the function of the *FAM184A* gene, microarray expression data from the USCS Human Genome Browser has shown it may be expressed in the thalamus, medulla oblongata, and fetal brain (34). The *IGSF9B* gene is believed to have very high expression in the brain, specifically the amygdala, corpus callosum, cerebellum, and caudate nucleus (35). However, none of the segments that overlapped in these gene regions showed significant association with schizophrenia risk, nor did we identify any locus that likely harbors risk alleles.

Limitations

Several limitations may have contributed to the outcome of the results. First, the mating relationships and demography of the population is not entirely clear. Although the clans were believed to practice exogamy, close relationships may have lead to potential inbreeding that could have been influenced by inbreeding depression. Although the mechanism of inbreeding depression is not well understood, it is possible it could have had an impact on the data given the close relationships of the subjects. In addition, the sample size of the study may not have allowed us to see regions that were shared by only a few subjects. Risk alleles may exist throughout the population of Palau, but not heavily present in our sample. Finally, it is possible that highly penetrant recessive risk variants may not exist in the genome of this population. Reduced penetrance and variable expressivity may be playing roles in the expression of schizophrenia and related disorders that were inhibiting our identification of recessive risk alleles.

Future Studies

The results of our study show that the substantial portion for risk of psychosis in Palau can be contributed to neither a common recessive variant, nor a rare, highly penetrant variant.

Alternative mechanisms for the association between homozygous segments and risk for psychotic disorders may be plausible. For instance, it is possible that the homozygous regions contain many variants of modest additive impact on risk, which are joined by chance as a haplotype; and homozygosity of that haplotype doubles the risk. Future studies may be able to explore alternative inheritance patterns outside of highly penetrant recessive risk variants to explain the increased prevalence of schizophrenia with increased homozygosity.

Conclusions

Our findings were consistent with previous studies that have reported an increased risk for schizophrenia when close mating relationships are practiced, resulting in an increase of HBD regions (11, 17). Despite this consistency, we did not find that the homozygous segments in Palau trace to the kind of risk loci hypothesized to generate the results of other studies. It is possible that the Palau population may be different from other populations examined, or that the mechanisms that associate excess regions of homozygosity to risk for schizophrenia is one that is not yet understood. Alternative explanations for why increased regions of homozygosity increase risk for schizophrenia include the reduction of fitness that is associated with inbreeding (38). In addition, studies have shown that an increase risk for schizophrenia is associated with an increased number of HBD segments (11). Whatever the explanation, this study illustrates the complexity of discovering genes related to complex diseases through homozygosity mapping.

The complexity of schizophrenia as well as other psychiatric disorders continues to impede efforts in solving the genetic puzzle that encodes its development. In working to develop the field of psychiatric genetic counseling, it is important to make an attempt in understanding the genetics of these neuropsychological illnesses. Identifying causative genes that can eventually be tested for can provide new opportunities for patients and their families. A few

findings have already begun to bring about discussion of psychiatric illnesses in a clinical setting. For example, 25-30% of patients with a 22q11 deletion syndrome (DiGeorge syndrome, Velocardiofacial syndrome) have schizophrenic behaviors (19, 44). It is predicted that 0.5-2% of all schizophrenia patients may have a deletion of 22q11. If this finding is identified through routine chromosome analysis or SNP microarray testing, then a conversation about the development of schizophrenic features could take place.

Any discovery of the molecular etiology related to a diagnosis of schizophrenia could impact the health management of a patient. With any genetic condition, concerns regarding recurrence risk, family planning, and issues involving stigma are present. Understanding the genetics means understanding the inheritance pattern, which gives a better picture of who is at risk and whether or not the condition will reoccur in the family. It also provides options for reproduction including prenatal genetic diagnosis and preimplantation genetic diagnosis (PGD). It also may give patients and families a better sense of control as to how their diagnosis happened, and what they can do to better monitor it.

Continuing efforts to identify risk alleles for schizophrenia is crucial in order for the field of psychiatric genetic counseling to grow. All areas of genetic counseling, including prenatal, cancer, and pediatrics, began with the discovery of few genetic links to disease, and then expanded. Cancer genetics would not be what it is today without the discover of the BRCA genes. Diagnosis of Down syndrome would not be as sufficient today had it not been for the discovery of the extra 21st chromosome in affected patients. Although this study did not identify any risk loci for schizophrenia, the concept of fine-mapping displayed through this project may one day uncover region clearly associated with risk, and the foundation of psychiatric genetics and counseling can continue to build.

APPENDIX A

RELATIONSHIP MATRICES

Table 7. Genetic relationship matrix for subjects who are carriers of homozygous haplotypes in RAB3GAP2

	Affected subjects homozygous for *RAB3GAP2* haplotype													
Affected	**13881**	**20256**	**20283**	**20310**	**22355**	**22357**	**22382**	**22416**	**22420**	**22476**	**22518**	**22522**	**5000**	**60041**
13881	1	0.509	0.007	0.5	0.012	0.000001	0	0	0.0085	0.002	0	0.000001	0.000001	0.0085
20256		1	0.0039	0.49	0.0156	0	0	0	0.0031	0.0071	0.004	0	0.0031	0.00009
20283			1	0.000002	0.002	0.000001	0.016	0.0063	0.000001	0.127	0	0.0006	0.000001	0.0146
20310				1	0.000001	0.000003	0	0	0	0.0097	0.000001	0	0	0.004
22355					1	0.000001	0	0.000001	0	0.0029	0.000004	0.000004	0.0159	0
22357						1	0.015	0.000001	0	0	0	0	0	0
22382							1	0.011	0.0046	0.0012	0.0058	0.000001	0.0062	0.00027
22416								1	0	0.0026	0.000004	0.0066	0	0.0075
22420									1	0.0099	0.0013	0	0.0027	0.087
22476										1	0.0085	0	0	0.000001
22518											1	0.000001	0	0.0135
22522												1	0.0071	0
5000													1	0.000001
60041														1

Table 7. Continued

	Unaffected subjects homozygous for RAB3GAP2 haplotype							
Affected	**26063**	**60032**	**60033**	**60072**	**60085**	**60114**	**1034**	**20268**
13881	0.000003	0.0038	0.000003	0.0024	0	0.0088	0.000001	0
20256	0.000001	0	0.0005	0.019	0	0.0116	0	0
20283	0.000001	0.00167	0.000004	0.0153	0	0.0159	0	0.012
20310	0.0012	0.0071	0.016	0.00015	0	0.0052	0	0
22355	0.011	0.00031	0	0.000001	0.000004	0	0	0.0688
22357	0.0019	0	0.013	0.0085	0.000007	0	0.0043	0.000012
22382	0.0149	0.012	0.000001	0.0064	0.0145	0.000001	0.0379	0.0303
22416	0.000006	0.0079	0.0166	0.000001	0.0086	0	0.000001	0.0235
22420	0.00041	0	0.000001	0.000012	0.000002	0.00085	0.204	0.0024
22476	0.00284	0.0389	0.0526	0.0163	0.000001	0.0151	0.0117	0
22518	0.0123	0.0047	0.000079	0.0037	0	0.000001	0.0177	0
22522	0	0	0.0226	0.0161	0.000001	0.0085	0	0.0127
5000	0.0191	0.000001	0.0152	0	0.0098	0.0033	0.000001	0.112
60041	0.0079	0.0186	0	0.000001	0.000001	0	0.0688	0.0154
Unaffected								
26063	1	0.0633	0.0009	0.0139	0.000001	0.0015	0.0053	0.085
60032		1	0.0065	0.0146	0.0082	0.000001	0.0143	0.0051
60033			1	0.000004	0.0188	0.000001	0.0207	0.0164
60072				1	0	0.000001	0.0216	0
60085					1	0	0.00998	0.000499
60114						1	0.003	0.000002
1034							1	0.0106
20268								1

Table 8. Genetic relationship matrix for subjects who are carriers of homozygous haplotypes in MARK1

	Affected subjects homozygous for MARK1 haploptype									Unaffected subjects homozygous for *MARK1* haplotype		
Affected	**20280**	**2058**	**22361**	**22420**	**25012**	**25043**	**30063**	**4977**	**4981**	**26096**	**60068**	**3011**
20280	1	0.0119	0.004	0.000001	0.000002	0.000001	0.000001	0.0076	0.008	0.026	0.039	0
2058		1	0.002	0	0.006	0.0048	0.000001	0.046	0.065	0.0009	0.001	0.004
22361			1	0.000001	0	0.0079	0.0357	0.000001	0.000001	0.000001	0.0070	0.0322
22420				1	0.0002	0	0.0185	0	0	0.00009	0	0
25012					1	0.0037	0.000001	0.0077	0.013	0.005	0.000001	0.000001
25043						1	0.000001	0	0	0.0099	0	0.036
30063							1	0	0.000004	0.0046	0.0018	0.051
4977								1	0.557	0.013	0	0.000001
4981									1	0.0004	0	0
Unaffected												
26096										1	0.021	0.0005
60068											1	0
3011												1

Table 9. Genetic relationship matrix for subjects who are carriers of homozygous haplotypes in DISC1

	Affected subjects homozygous for *DISC1* haplotype							Unaffected subjects homozygous for *DISC1* haplotype	
Affected	**20269**	**20279**	**22354**	**22355**	**22518**	**60041**	**22536**	**22442**	**30019**
20269	1	0.000004	0.0186	0.003	0.0028	0	0.0086	0	0
20279		1	0	0	0	0.009	0.0166	0.0027	0.006
22354			1	0.025	0	0.010	0.0048	0	0.016
22355				1	0.000004	0	0	0.0058	0
22518					1	0.013	0.000004	0	0
60041						1	0.209	0.0011	0.015
22536							1	0.0021	0
Unaffected									
22442								1	0.0016
30019									1

Table 10. Genetic relationship matrix for subjects who are carriers of homozygous haplotypes in FAM184A

	Affected subjects homozygous for *FAM184A* haplotype							Unaffected subjects homozygous for *FAM184A* haplotype
Affected	**17618**	**20304**	**20306**	**22357**	**22386**	**22392**	**4977**	**24994**
17618	1	0.000000	0.007508	0.005252	0.000001	0.007559	0.009239	0.000006
20304		1	0.000001	0.000002	0.004128	0.008642	0.011500	0.000000
20306			1	0.003107	0.031520	0.000002	0.000000	0.006025
22357				1	0.000182	0.000001	0.000001	0.000001
22386					1	0.000001	0.019165	0.000004
22392						1	0.000000	0.000001
4977							1	0.000001
Unaffected								
24994								1

Table 11. Genetic relationship matrix for subjects who are carriers of homozygous haplotypes in IGSF9B

	Affected subjects homozygous for IGSF9B haplotype													
Affected	**19511**	**20248**	**20253**	**20267**	**20269**	**20309**	**20323**	**22368**	**22432**	**25012**	**60015**	**60027**	**60047**	**20308**
19511	1	0.000003	0.002749	0.000004	0.000001	0.319471	0.004834	0.007868	0.006417	0.01269	0.00001	0.004595	0.003732	0.055844
20248		1	0.003888	0.000000	0.000000	0.000000	0.000000	0.409338	0.000001	0.010057	0.000000	0.000001	0.000995	0.000002
20253			1	0.000004	0.001119	0.010444	0.000000	0.000001	0.000001	0.009480	0.000002	0.012884	0.001539	0.005695
20267				1	0.000001	0.000000	0.000000	0.014302	0.000003	0.000834	0.000480	0.015920	0.426976	0.007413
20269					1	0.000000	0.021727	0.000001	0.000006	0.001338	0.000001	0.002884	0.000975	0.008583
20309						1	0.003105	0.001889	0.000000	0.008099	0.004081	0.000001	0.000000	0.032079
20323							1	0.009329	0.001354	0.000188	0.000001	0.001915	0.000000	0.000001
22368								1	0.008936	0.000000	0.000000	0.003695	0.008104	0.000449
22432									1	0.009401	0.000000	0.014157	0.000001	0.002326
25012										1	0.000001	0.000002	0.013627	0.000000
60015											1	0.000576	0.000000	0.000001
60027												1	0.002056	0.006798
60047													1	0.011744
20308														1

Table 11. Continued

	Unaffected subjects homozygous for IGSF9B haplotype						
Affected	**25000**	**26063**	**26135**	**60004**	**60008**	**60032**	**60080**
19511	0.005192	0.000000	0.015889	0.087332	0.000001	0.002361	0.000008
20248	0.000001	0.021340	0.001075	0.004235	0.017008	0.000002	0.000000
20253	0.000000	0.011792	0.010291	0.006911	0.000002	0.000001	0.000001
20267	0.001798	0.019823	0.044625	0.002305	0.000000	0.003560	0.004506
20269	0.000000	0.000000	0.000000	0.000001	0.016779	0.000789	0.000000
20309	0.012613	0.003395	0.010645	0.044668	0.000001	0.000000	0.008291
20323	0.000000	0.006682	0.000004	0.008460	0.000000	0.000004	0.008744
22368	0.000000	0.018422	0.000974	0.000001	0.004808	0.000509	0.000007
22432	0.004383	0.063062	0.014521	0.003708	0.008141	0.082299	0.037833
25012	0.000000	0.011510	0.007338	0.000004	0.000000	0.025294	0.000874
60015	0.000000	0.000001	0.004763	0.002771	0.004211	0.000001	0.000001
60027	0.004597	0.000000	0.000000	0.002907	0.005236	0.010568	0.002771
60047	0.000000	0.023770	0.027584	0.000001	0.000001	0.000178	0.001561
20308	0.014241	0.001996	0.004802	0.053688	0.000002	0.000004	0.003645
Unaffected							
25000	1	0.000000	0.017069	0.000079	0.000001	0.000092	0.000004
26063		1	0.007281	0.000001	0.000001	0.063297	0.001913
26135			1	0.018914	0.011637	0.003062	0.006551
60004				1	0.008243	0.005106	0.000002
60008					1	0.006010	0.006974
60032						1	0.020900
60080							1

APPENDIX B

IRB APPROVAL FORM

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: Bernard Devlin, PhD
From: Christopher Ryan, PhD, Vice Chair
Date: 8/15/2012
IRB#: REN12080113 / PRO08040446
Subject: The Genetics of Schizoprhrenia in Oceanic Palau

Your renewal for the above referenced research study has received expedited review and approval from the Institutional Review Board under: **This approval is for analysis of data only.**

45 CFR 46.110.(7) characteristics/behaviors

Please note the following information:

Approval Date: 8/15/2012
Expiration Date: 8/14/2013

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.

BIBLIOGRAPHY

1. American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders, fourth edition. Washington, DC. 273-315.

2. Bertelsen, A. (2002). Schizophrenia and related disorders: experience with current diagnostic systems. *Psychopathology*, **35**, 89-93.

3. Devlin, B., Bacanu, S-A., Roeder, K., Reimherr, F., Wender, P., Galke, B., *et al*. (2002). Genome-wide multipoint linkage analyses of multiplex schizophrenia pedigrees from the oceanic nation of Palau. *Molecular Psychiatry*, **7**, 689-694.

4. National Institute of Mental Health. (2009). Schizophrenia. *U.S. Department of Health and Human Services*, 1-17.

5. Flaum, M., Andreasen, N. C. (1991). Diagnostic criteria for schizophrenia and related disorders: options for DSM-IV. *Schizophrenia Bulletin*, **17**(1), 133-142.

6. Tsuang, M. (2000). Schizophrenia: genes and environment. *Biol Psychiatry*, **47**, 210-220.

7. Tsuang, M. T., Gilbertson, M. W., Faraone, S. V. (1991). The genetics of schizophrenia: current knowledge and future directions. *Schizophrenia Reseach*, **4**, 157-171.

8. Bourgain, C., Genin, E., Quesneville, H., Clerget-Darpoux, F., (2000). Search for multifactorial disease susceptibility genes in founder populations. *Ann. Hum. Genet*, 64, 255-265.

9. Myles-Worsley, M., Tiobech, J., Blailes, F., Middletone, F. A., Vinogradov, S., Byerley, W., et al. (2010). Familial transmission of schizophrenia in Palau: a 20-year genetic epidemiological study in three generations. *Amer. J. Med. Genet*, 247-254.

10. Lu, C. (2011). Homozygosity by descent and risk for schizophrenia in Palau. 1-18.

11. Keller, M. C., Simonson, M. A., Ripke, S., Neale, B. M., Gejman, P. V., Howrigan, D. P., et al. (2012). Runs of homozygosity implicate autozygosity as a schizophrenia risk factor. *PLoS Genetics*, **8**(4), 1-11.

12. Lencz, T., Lambert, C., DeRosse, P., Budrick, K. E., Morgan, V., Kane, J. M., et al. (2007). Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *PNAS*, **104**(50), 19942-19947.

13. Heutink, P., Oostra, B. A. (2002). Gene finding in genetically isolated populations. *Human Molecular Genetics*, **11**(20), 2507-2515.

14. Browning, S. R., Browning, B. L. (2010). High-resolution detection of identity by descent in unrelated individuals. *The American Journal of Human Genetics*, **86**, 526-539.

15. Myles-Worsley, M., Coon, H., Tiobech, J., Collier, J., Dale, P., Wender, P., et al. (1990). Genetic epidemiological study of schizophrenia in Paulau, Micronesia: prevalence and familiality. *Am. J. Med. Genet*, **88**, 4-10.

16. Devlin, B., Bacanu, S-A., Roeder, K., Reimherr, F., Wender, P., Galke, B., et al. (2002). Genome-wide multipoint linkage analyses of multiplex schizophrenia pedigrees from the oceanic nation of Palau. *Molecular Psychiatry*, **7**, 689-694.

17. Mansour, H., Fathi, W., Lambertus, K., Wood, J., Chowdari, K., Watson, A., et al. (2010). Consanguinity and increased risk for schizophrenia in Egypt. *Schizophrenia Research*, **120**, 108-112.

18. Melhem, N., Middleton, F., McFadden, K., Klei, L., Faraone, S. V., Vinogradov, S., et al. (2011). Copy number variants for schizophrenia and related psychotic disorders in oceanic Palau: risk and transmission in extended pedigrees. *Biol Psychiatry*, **70**, 1115-1121.

19. Hodgkinson, K. A., Murphy, J., O'Neill, S., Brzustowics, L., Bassett, A. S. (2001). Genetic counseling for schizophrenia in the era of molecular genetics. *Can J Psychiatry*, **46**(2), 123-130.

20. Rice, D. P., Miller, L. S. (1998). Health economics and cost implications of anxiety and other mental disorders in the United States. *Br. J. Psychiatry*, **173**, 4-9.

21. Austin, J. C., Honer, W. G. (2007). The genomic era and serious mental illness: a potential application for psychiatric genetic counseling. *Psychiatric Services*, **58**(2), 254-261.

22. Thonton, T., McPeek, M. S. (2007). Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *Am. J. Hum. Genet.* **81**, 321-337.

23. Online Mendelian Inheritance in Man, OMIM. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). Nov 20, 2009. <http://omim.org/entry/610836>.

24. Online Mendelian Inheritance in Man, OMIM. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). Oct 31, 2011. <http://omim.org/entry/212720>.

25. Online Mendelian Inheritance in Man, OMIM. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD). Feb 21, 2012. <http://omim.org/entry/605210>.

26. Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J. et al. (2009). Finding the missing heritability of complex diseases. *Nature*, **461**, 747-753.

27. Devlin, B., Roeder, K., Otto, C., Tiobech, S., Byerley, W. (2001). Genome-wide distribution of linkage disequilibrium in the population of Palau and its implications for gene flow in Remote Oceania. *Hum Genet*, **108**, 521-528.

28. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., et al. (2007). PLINK: a toolset for whole-genome association and population-linkage analysis. *Amer J Hum Genet*, **81**, <http://pngu.mgh.harvard.edu/purcell/plink>.

29. World Health Organization. (1996). Schizophrenia and public health. *Division of Mental Health and Prevention of Substance Abuse*. WHO/MSA/NAM/97.6.

30. Weir, Bruce. (1943). *Genetic Data Analysis*. Sinauer Associates, Inc, Sunderland, MA.

31. Slater, E., Cowie, V. (1971). *The Genetics of Mental Disorders*. Oxford University Press, London.

32. Kety, S. S., Rosenthal, D., Wender, P. H., Schulsinger, F., Jacobsen, B. (1976). Mental illness in the biological and adoptive families of adopted individuals who have become schizophrenic. *Behavior Genetics*, **6**(3), 219-225.

33. Stefansson, H., Rujescu, D., Cichon, S., Pietilainen, O. P., Ingason, A., Steinberg, S., et al. (2008). Large recurrent microdeletions associated with schizophrenia. *Nature*, **455**, 237-241.

34. Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., Haussler, D. (2002). The human genome browser at UCSC. *Genome Res*. **12** (6), 996-1006.

35. Online Mendelian Inheritance in Man. OMIM. McKusick-Nathans Institute of Genomic Medicine, Johns Hopkins University (Baltimore, MD). Feb 24, 2011. <http://omim.org/entry/613773>.

36. Online Mendelian Inheritance in Man. OMIM. McKusick-Nathans Institute of Genomic Medicine, Johns Hopkins Univerrsity (Baltimore, MD). Sept 14, 2011. <http://omim.org/entry/609275>.

37. Online Mendelian Inheritance in Man. OMIM. McKusick-Nathans Institute of Genomic Medicine, Johns Hopkins University (Baltimore, MD). Nov 20, 2009. <http://omim.org/entry/606511>.

38. Charlesworth, D., Willis, J. H. (2009). The genetics of inbreeding depression. *Nature: Reviews*. **10**, 783-796.

39. Klei, L., Bacanu, S., Myles-Worsley, M., Galke, B., Xie, W., Tiobech, J., et al. (2005). Linkage analysis of completely ascertained sample of familial schizophrenics and bipolar from Palau, Micronesia. *Hum Genet*. **117**, 349-356.

40. Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet*. 1-9.

41. O'Donovan, M. C., Craddock, N., Norton, N., Williams, H., Peirce, T., Moskvina, V., et al. (2008). Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nature Genetics*. **40**(9), 1053-1055.

42. Xu, B., Ionita-Laza, I., Roos, J. L., Boone, B., Woodrick, S., Sun, Y., et al. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nature Genetics*. **44**(12), 1365-1369.

43. Ripke, S., Sanders, A. R., Kendler, K. S., Levinson, D. F., Sklar, P., Holmans, P. A. et al. (2011). Genome-wide association study identifies five new schizophrenia loci. *Nature Genetics*. **43**(10), 969-974.

44. International Schizophrenia Consortium. (2008). Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*. **455**, 237-241.