FINE-MAPPING AND MUTATION IDENTIFICATION FOR ARVC (ARRHYTHMOGENIC RIGHT VENTRICULAR CARDIOMYOPATHY) AT THE ARVD5 LOCUS (3p25) IN THE NEWFOUNDLAND POPULATION

VANESSA FRENCH

Fine-mapping and mutation identification for ARVC (Arrhythmogenic Right Ventricular Cardiomyopathy) at the ARVD5 locus (3p25) in the Newfoundland population.

By

Vanessa French, BSc

A thesis submitted to the School of Graduate
Studies for partial fulfillment of the
requirements for the degree of
Master of Science

Discipline of Genetics
Faculty of Medicine
Memorial University of Newfoundland
St. John's, Newfoundland

November 2008

ABSTRACT

Cardiomyopathy is a disease that affects the muscle of the heart. It has a variety of causes, symptoms, and treatments. As the disease progresses, the heart becomes weaker and less able to pump blood through the body. There are four major groups of cardiomyopathy, which include Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC), Dilated Cardiomyopathy (DCM), Hypertrophic Cardiomyopathy (HCM), and Restrictive Cardiomyopathy (RCM).

Mutations causing ARVC have been previously found in seven genes. Five of these *(PKP2*, *DSP*, *DSC2*, *JUP*, *DSG2*) code for desmosomal proteins and are functionally important for cell-to-cell adhesion. A sixth gene, *RYR2,* is a phenocopy involved in calcium release mechanisms. A seventh gene, *TGFβ3,* is implicated in rare cases of ARVC. One of the twelve loci implicated in ARVC is designated Arrhythmogenic Right Ventricular Dysplasia Type 5 (ARVD5). Inheritance is in an autosomal-dominant pattern. It, like desmosomal ARVC, is characterized predominantly by the replacement of cardiomyocytes with fatty and fibrous tissue, which can increase the risk of sudden cardiac death (SCD). In 1998, linkage analysis and a genome wide scan, using a family from Newfoundland, identified an autosomal-dominant form of ARVC mapping to a 9.93 Mb region on chromosome 3p25.

The primary objective of this study was to reduce the linked candidate region, by identifying recombinations and/or a consistent disease-associated haplotype in additional ARVD5-linked Newfoundland families. The secondary objective was to identify the disease-causing gene.

Genomic DNA was collected from obligate carriers and affected members from 15 Newfoundland families and was used to construct high-resolution haplotypes. Fine mapping reduced the candidate region from 9.93 Mb to 2.36 Mb. A common haplotype co-segregated with the disease suggesting a common founder in these 14 Newfoundland families and a common mutation causing the disease. Direct sequencing of twenty positional candidate genes identified a novel missense mutation (c.1073 C→T, S358L), in the *TMEM43* gene, which was carried by all ARVC patients and absent from 322 population control chromosomes. *TMEM43* is a novel causative gene for ARVC at the ARVD5 locus. Presymptomatic diagnosis of at-risk individuals can be carried out due to this mutation discovery.

ACKNOWLEDGEMENTS

Completing a thesis is a challenge and thanking all those who contributed to it is an even greater one. Due to the nature of my study, many people were involved, providing both molecular and clinical expertise. I want to thank all of them, including those not mentioned here by name. A special thank you to my supervisor, Dr. Terry-Lynn Young, for inspiring my interest in ARVC research and providing invaluable guidance to our research teams.

I would like to express my sincere thanks to Dante Galutira for his unwavering patience during my initial training period. Ingrid Pardoe was a lab manager extraordinaire whose many lab duties included DNA bank management for my project. Both Dante and Ingrid worked with me during the first half of the genotype data analysis process. As our Sequencing Team leader, Nancy Merner's energy and passion for this research was remarkable. Both Nancy and Dr. Annika Haywood's perseverance was key as we worked through the sequencing phase of this project. Clinical ascertainment, counseling and treatment of patients were carried out by a full clinical team, including - but not exclusive to - cardiologist Dr. Sean Connors and genetic counselors Kathy Hodgkinson and Sara MacKay. I also wish to thank all the doctors, who have referred patients to this study, and the patients who have given their trust, time and samples to us. Without their generous involvement, this research would not have been possible.

I would like to thank all my family and friends. In particular my mom, Nancy, and my dog, Edie. Without their support and understanding I would have cracked up long ago. If I had more time to write another thesis I would express how grateful I am for everything; whether it was food donations, doing my laundry, housecleaning (!), hugs and wet kisses, cuddles, and relaxing walks surrounded by nature (I'll let the readers decipher who did what). I cannot adequately express how thankful I am. xo

Finally, I close with personal gratitude for having experienced a deep and energizing passion for this very special research project on ARVC and for the privilege of having helped make a difference in the lives of many people.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

D5 Family 185

D6 Family 273

D7 Family 453

D8 Family 581

D9 Family 840

D10 Family 853

D11 Family 864

D12 Family 932

D13 Family 964

D14 Family 977

D15 Family 1139

LIST OF PEDIGREE SYMBOLS

Female, unaffected or clinical status unknown

Female, affected with primary affection status

Male, unaffected or clinical status unknown

Male, affected with primary affection status

Female obligate carrier

Male obligate carrier

Female, affected with secondary affection status

Male, affected with secondary affection status

Mating

Consanguineous mating

Deceased

Proband

GLOSSARY

Adipogenesis, fat formation

Genetic distance, a measurement of the amount of evolutionary divergence between individuals of the same or different species. It is measured by counting the number of allelic substitutions per locus that have cropped up in each set of individuals

Microsatellite markers, a short (up to several hundred base pairs) segment of DNA that consists of multiple tandem repeats of a two or three base pair sequence

Nonepidermolytic palmoplantar keratoderma, thickening of the palms and soles as a result of excessive keratin formation

Non-ischemic ventricular arrhythmias, abnormal pumping of the heart

Pre-syncope, light headedness

Syncope, fainting

Ventricular tachycardia, abnormal rhythm

CHAPTER 1: LITERATURE REVIEW

1.1 ARVC

1.1.1 Clinical Manifestations of ARVC

Diseases of the cardiac muscle can be classified into four major groups: arrhythmogenic right ventricular cardiomyopathy (ARVC), dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), and restrictive cardiomyopathy (RCM). Analysis of familial cardiomyopathies frequently confirms Mendelian patterns of segregation (Marian et al., 1992; Mogensen et al., 2003; Song et al., 2007). ARVC is familial rather than sporadic in the majority of cases (Sen-Chowdhry et al., 2007). It has been reported that 20-65% of DCM cases are familial, with a predominance of autosomal-dominant inheritance pattern (Arimura et al., 2004; Bienengraeber et al., 2004; Gerull et al., 2002; Kamisago et al., 2000; Murphy et al., 2004; Olson, Kishimoto, Whitby, & Michels, 2001; Olson et al., 2002; Schmitt et al., 2003; Schonberger et al., 2000, Mestroni et al., 1999). Familial HCM occurs as an autosomal dominant disease in approximately 50% of cases and is a genetically heterogenous disease, in that it can be caused by genetic defects at more than one locus (Bonne, Carrier, Richard, Hainque, & Schwartz, 1998). Approximately 30% of children with RCM have a family history of cardiomyopathy (Chen, Balfour, & Jureidini, 2001; Denfield et al., 1997; Gewillig, Mertens, Moerman, & Dumoulin, 1996; Lewis, 1992)

ARVC is a hereditary heart muscle disorder with clinical manifestations of heart palpitations, pre-syncope, syncope and sudden death. Most cases of ARVC are inherited in an autosomal dominant fashion, with only a few documented cases of autosomal recessive mode of inheritance – causing Carvajal syndrome and Naxos disease (G. Thiene et al., 2007). ARVC is thought to be familial in 30–70% of cases (Hamid et al., 2002). It is clinically characterized by non-ischemic ventricular arrhythmias. These arrhythmias arise predominantly from the right ventricle and are potentially lethal. Pathologically, ARVC is diagnosed with an endomyocardial biopsy, or upon postmortem examination, and is characterized by progressive dystrophy of the myocardium with fibro-fatty replacement. The results of this myocardial degradation include ventricular arrhythmias, structural and functional abnormalities predominantly of the right ventricle, electrocardiographic depolarization/repolarisation changes, and sudden death (Sen-Chowdhry, Lowe, Sporton, & McKenna, 2004).

Clinical diagnosis of ARVC has proved to be difficult in the absence of a pathognomonic feature. In 1994, an international task force was established to address this issue by creating a set of criteria for ARVC diagnosis. Disease features were designated *minor* or *major* criteria (Table 1). It was agreed that a positive diagnosis could be confirmed only in the presence of the following criteria combinations: two major criteria; one major plus two minor criteria; or, four minor criteria across the six diagnostic fields (McKenna et al., 1994). A major or minor criterion is counted only once in each field and minor criteria are counted only in the absence of a major criterion within a field.

Table 1: Task Force Criteria for the Diagnosis of Right Ventricular Dysplasia in Probands

Family history
Major: familial disease confirmed at necropsy or surgery.
Minor: family history of premature sudden death (<35 years of age) due to suspected right ventricular dysplasia; family history (clinical diagnosis based on present criteria).
ECG depolarization/conduction abnormalities
Major: epsilon waves or localized prolongation (>110 ms) of QRS complex in right precordial leads (V1–V3).
Minor: late potentials on signal-averaged ECG.
ECG repolarization abnormalities
Minor: inverted T waves in right precordial leads (V2 and V3) in persons >12 years of age and in the absence of right bundle branch block.
Arrhythmias
Minor: sustained or nonsustained left bundle branch block-type ventricular tachycardia documented on ECG or Holter monitoring or during exercise testing; frequent ventricular extrasystoles (>1000/24 hours on Holter monitoring).
Global or regional dysfunction and structural alterations
Major: severe dilatation and reduction of right ventricular ejection fraction with no or mild left ventricular involvement; localized right ventricular aneurysms (akinetic or dyskinetic areas with diastolic bulgings); severe segmental dilatation of right ventricle.
Minor: mild global right ventricular dilatation or ejection fraction reduction with normal left ventricle; mild segmental dilatation of right ventricle; regional right ventricular hypokinesia.
Tissue characteristics of walls
Major: fibrofatty replacement of myocardium on endomyocardial biopsy.

From: (McKenna et al., 1994)

ARVC linked to 3p25 manifests itself predominantly after adolescence. This late onset may be problematic when assigning affection status to ARVC family members. Caution must be used when interpreting the affection status of individuals classified as "unaffected". Males can experience first symptoms, including SCD, by the second decade of life. Females generally have an onset in their 3rd to 5th decade of life. This age-dependant penetrance - in addition to diagnostic difficulties - results in under-recognition of mutation carriers until symptoms present.

Penetrance is the probability that a particular phenotype is expressed in an individual that carries the disease susceptibility genotype. Penetrance can be age related, meaning that the trait is not expressed in most carriers at birth but occurs with increased frequency as the carriers get older (Haines & Pericak-Vance, 2006). Assuming that a single disease locus is bi-allelic, with a dominant "A" allele and a recessive "a" allele, then there exists up to three combinations of genotypes carried by individuals in the population – AA, Aa, and aa. Individuals who carry the dominantly inherited susceptibility disease allele A, either AA or Aa genotype, have the same non-zero probability of disease onset when there is incomplete penetrance - less than 100%. For example, gastric cancer, a dominant susceptibility disease, was investigated in four families originating from the southeastern coast of Newfoundland. It was reported that the penetrance rate of gastric cancer was 40% for males and 63% for females by the age of 75 years. This means that 60% of males and 37% of patients who carry the known mutations will not show clinical signs of the disease at age 75 years, respectively (Kaurah et al., 2007). Asymptomatic disease-gene carriers can appear to be unaffected due to incomplete penetrance. This can lead to under-recognition of the disease in the population. ARVC populations around the world have shown variable penetrance (Moric-Janiszewska & Markiewicz-Loskot, 2007), that may lead to an under-recognized disease and a prevalence rate that is difficult to accurately calculate.

The prevalence of ARVC in the general population is difficult to determine. This is due to either a clustering of the disease in some geographic areas or the underdiagnosis and

misdiagnosis of patients, both pathologically or clinically (Corrado et al., 2000; G. Thiene et al., 1997). For example, in a recent study of 89 probands diagnosed with ARVC, reevaluation using the Task Force Criteria (Table 1) found that only 24 (27%) met the diagnosis criteria. Replication of the initial magnetic resonance imaging (MRI) testing could not confirm the ARVC diagnosis in 46 of the 89 patients (Bomma et al., 2004). In the United States, approximately 1:5,500, or 0.02% of the population, suffer from some form of cardiomyopathy at any given time (US Census Bureau, Population Estimates, 2004). In the Veneto region of Italy, the prevalence of ARVC has been estimated from 1:2,000 to 1:5,000, accounting for 20% of sudden cardiac deaths occurring below the age of 35 years (Nava et al., 1988). The prevalence of ARVC in the Newfoundland population is currently unknown.

1.1.2 Management and Therapy

The primary treatment goal in ARVC is prevention of sudden cardiac death (SCD). It is of the utmost importance that individuals at risk of ARVC be identified so that protective measures can be offered. The implantable cardioverter-defibrillator (ICD) has proved to be the best protection against SCD. The ICD constantly monitors the heart's rhythm and when ventricular tachycardia (VT) is detected, the ICD attempts to restore normal sinus rhythm by delivering a high-voltage synchonizing impulse to the heart (Hodgkinson et al., 2005). Antiarrhythmic drugs, such as sodium channel blockers, β-blockers, sotalol, amiodarone and verapamil are also effective. Catheter ablation has also been successively applied to ARVC patients (Wichter et al., 2005). Although its beneficial

effects are short term, the survival outcome is quite good (40% freedom from recurrence at 3 years) (G. Thiene et al., 2007).

1.1.3 Genetics and Cellular Mechanism of ARVC

Twelve loci causing ARVC have been mapped and mutations have been previously found in seven genes. Their gene products code for either desmosomal proteins (plakophilin-2 (*PKP2*), desmoplakin (*DSP*), desmocolin-2 (*DSC2*), plakoglobin (*JUP*), desmoglein-2 (*DSG2*), cardiac ryanodine receptor 2 (*RYR2* – a phenocopy) or transforming growth factor beta-3 (*TGFβ3*) (Table 2).

Early attempts to identify disease-causing genes in ARVC involved a recessive variant shown to cause Naxos disease, characterized by nonepidermolytic palmoplantar keratoderma and woolly hair. This obvious external phenotype allowed for easy identification of affected individuals. In 1998, linkage analysis mapped the genetic locus to 17q21 (Coonar et al., 1998) and two years later a candidate gene approach study reported a two base pair (bp) deletion in the plakoglobin gene (*JUP*) as the cause of Naxos disease. This mutation resulted in a frameshift and premature termination of the protein (McKoy et al., 2000). Several years later, the gene was screened in a German family with ARVC but no cutaneous abnormalities. A dominant mutation in the gene was identified and further analysis suggested that the mutation may increase plakoglobin turnover via proteasomal degradation (Asimaki et al., 2007). Plakoglobin is a major constituent of cell adhesion junctions and its role in Naxos disease led to a search for

other ARVC genes with related functions. Several months later, a linkage analysis study reported a deletion in desmoplakin (*DSP*) as the causative mutation in the recessive Carvajal syndrome, characterized by generalized striate keratoderma, woolly hair and a dilated left ventricular cardiomyopathy (Norgett et al., 2000). Desmoplakin is an obligate component of functional desmosomes that anchors intermediate filaments to desmosomal plaques. Bioinformatic analysis anticipated that the mutation would cause a truncation of the C-terminal domain of the protein. In 2003, a linkage analysis and candidate gene approach was used to isolate a missense mutation in this *DSP* region in an Arab family with recessive ARVC, woolly hair, and a pemphigous-like skin disorder (Alcalai et al., 2003).

Three years after the first gene causing an autosomal recessive form of ARVC was identified (Naxos disease), the first gene causing an autosomal dominant form of ARVC was discovered - ryanodine receptor 2 (*RYR2*). This linkage analysis and candidate gene study involved four independent families from north-east Italy. The protein product of *RYR2*, composed of a tetramer of the ryanodine receptor proteins and a tetramer of FK506 binding protein 1B proteins, was located in the smooth sarcoplasmic reticulum and mediated calcium release for electroanatomical coupling (Tiso et al., 2001). *RYR2* mutations are also documented to cause catecholaminergic polymorphic ventricular tachycardia (CPVT). This disease has a phenotypically different rhythm, compared to ARVC, which is caused by calcium leaks in the sarcoplasmic reticulum (Liu, Rizzi, Boveri, & Priori, 2008; Tateishi et al., 2008). Thus, *RYR2* may cause a phenocopy of

ARVC due to sampling error of myocardial tissue at pathology (Awad, Calkins, Judge, & Medscape, 2008). Research, utilizing linkage analysis and screening of candidate genes, on an Italian family with the typical right ventricular phenotype, identified a missense mutation in the desmoplakin gene (Rampazzo et al., 2002). This mutation resulted in an amino acid substitution modifying a putative phosphorylation site in the N-terminal plakoglobin binding domain. Three years later, linkage analysis and a subsequent candidate gene screening reported that a single adenine insertion in the N-terminal of desmoplakin had been isolated in a family with typical arrhythmogenic left ventricular cardiomyopathy. This insertion is predicted to introduce a premature stop codon, resulting in truncation of the rod and C-terminal domains (Norman et al., 2005). In 2004, screening of the candidate gene plakophilin-2 (*PKP2*), an essential armadillo-repeat protein of the cardiac desmosome, in 120 unrelated ARVC diagnosed probands revealed several heterozygous mutations. *PKP2* protein was localized to cell desmosomes and nuclei, and participated in linking cadherins to intermediate filaments in the cytoskeleton (Gerull et al., 2004). The vast majority of mutations reported in the *PKP2* gene are heterozygous and result in missense, nonsense and frameshift mutations. However, a recessively inherited novel cryptic splice mutation has been reported (Awad et al., 2006). The candidate gene Transforming Growth Factor Beta-3 gene (*TGFβ*) was screened in 30 ARVC probands resulting in the identification of two mutations - a nucleotide substitution (c.-36G>A) in 5' UTR and a mutation (c.1723C>T) in the 3' UTR. The gene defect may have accounted for an increased tendency for extracellular matrix production and adipogenesis (Beffagna et al., 2005). In 2006, nine mutations in the candidate gene

desmoglein-2 (*DSG2*), the only desmoglein isoform expressed in cardiac myocytes and a calcium-binding transmembrane glycoprotein component of desmosomes, were identified in unrelated ARVC probands of Italian descent. Seven of the nine mutations were located in the extracellular amino terminal domain, which directly participated in the adhesive interaction (Pilichou et al., 2006). In the same year, the candidate gene desmocollin-2 (*DSC2*), a gene coding for a desmosomal cadherin, was screened in 77 ARVC probands. Two mutations, a deletion and an insertion, were identified. Both mutations resulted in frameshifts and premature truncation of the desmocollin-2 protein (Syrris, Ward, Evans et al., 2006). Another study investigated 88 unrelated patients and found a heterozygous splice-acceptor–site mutation in the *DSC2* gene. This resulted in a the use of a cryptic splice-acceptor site and the creation of a downstream premature termination codon (Heuser et al., 2006). The *DSC2* was also screened in 54 patients of Italian descent. Two heterozygous mutations were found in two probands and in four family members. These mutations map to the N-terminal region, relevant to adhesive interactions(Beffagna et al., 2007).

The cellular components linked to ARVC are summarized in Figure 1B, C, D. Proposed cellular mechanisms of desmosomal gene mutations and ARVC include: cell death and regional fibrosis resulting from the loss of myocyte adhesion (Asimaki et al., 2007; Lahtinen et al., 2008); adipogenesis by the suppression of the canonical Wnt/β-catenin signaling, due to the translocation of plakoglobin from the plasma membrane to the nucleus (Garcia-Gras et al., 2006); disruption of the normal localization and conductivity

of the gap junction protein, connexin 43, due to decreased expression of plakophilin-2 in cardiac cells (Oxford et al., 2007); and, an increase in arrhythmias by abnormal calcium storage mediated by a phenocopy histidine-rich calcium-binding protein (Hofmann et al., 1989; Lee, Kang, Kim, & Park, 2001). The functions of non-desmosomal gene mutations have also been studied and give insights to the cellular mechanisms. *TGFβ3* stimulates the production of components of the extracellular matrix. It can also down-regulate the synthesis of proteinases which degrade the matrix, often leading to excessive accumulation of connective tissue(Augusciak-Duma & Sieron, 2008; Sen-Chowdhry, Syrris, & McKenna, 2007), while an abnormal RYR2 protein can result in defective calcium channel function (Marx et al., 2000). To date, the disease-causing gene mutations for five of the twelve ARVC linked loci have yet to be identified, one of which is the ARVD5 locus (Table 2). Based on previously described ARVC disease-causing genes, functions of candidate genes at the ARVD5 locus may be important for cell-to-cell adhesion and conduction.

Table 2: Summary of Linked Loci and Genes Identified in AVRC/D

Designation	Gene (Symbol)	Locus	Mode of Inheritance	Type of Reported Mutations	Associated Phenotype	References
ARVD1	Transforming Growth Factor Beta-3 (TGFβ3)	14q23-q24	AD	Non-coding (5' untranslated region)	ARVC	(Beffagna et al., 2005; Nattel & Schott, 2005; Sen-Chowdhry, Syrris, & McKenna, 2005; Yang et al., 2006)
ARVD2	Cardiac Ryanodine Receptor 2 (RYR2)	1q41.2-q43	AD	Missense	ARVC	(Sen-Chowdhry et al., 2005; Tiso et al., 2001; Yang et al., 2006)
ARVD3	Unknown	14q12-q22	AD	Unknown	ARVC	(Severini et al., 1996)
ARVD4	Unknown	2q32.1-q32.3	AD	Unknown	ARVC	(Kirsch, Weinstock, Magid, Levin, & Gold, 1993; Rampazzo et al., 1997)
ARVD5	Unknown	3p25	AD	Unknown	ARVC	(Ahmad et al., 1998)
ARVD6	Unknown	10p14-p12	AD	Unknown	ARVC	(Li et al., 2000; Li, Gonzalez, Bachinski, & Roberts, 2000)
ARVD7	Unknown	10q22.3	AD	Unknown	ARVC	(Melberg et al., 1999)
ARVD8	Desmoplakin (DSP)	6p24	AD	Various	ARVC	(Corrado et al., 1997; Lindstrom, Nylander, Larsson, & Wranne, 2005; Nava et al., 1988; Richardson et al., 1996; G. Thiene et al., 1988)
			AR	Missense	ARVC, skin disorder, woolly hair	(Sen-Chowdhry et al., 2007)
			AR	Nonsense	ARVC, skin disorder, woolly hair	(Norman et al., 2005)
			AR	Deletion	Carvajal syndrome	(McKoy et al., 2000)
ARVD9	Plakophilin (PKP2)	12p11	AD	Various	ARVC	(Alcalai, Metzger, Rosenheck, Meiner, & Chajek-Shaul, 2003; Basso et al., 2006; Bauce et al., 2005; Corrado et al., 1997; Norgett et al., 2000; Norman et al., 2005; Rampazzo et al., 2002;

						Uzumcu et al., 2006; Yang et al., 2006)
			AR	Cryptic splice site	ARVC	(Alcalai et al., 2003; Basso et al., 2006; Bauce et al., 2005; Corrado et al., 1997; Norgett et al., 2000; Norman et al., 2005; Rampazzo et al., 2002; Uzumcu et al., 2006; Yang et al., 2006)
ARVD10	Desmoglein (DSG2)	18q12	AD	Various	ARVC	(Corrado et al., 1997; Gerull et al., 2004; Syrris, Ward, Asimaki et al., 2006; van Tintelen et al., 2006)
ARVD11	Desmocollin (DSC2)	18q12	AD	Deletion, insertion, splice site	ARVC	(Antoniades et al., 2006; Dalal et al., 2006)
ARVD12	Plakoglobin (JUP)	17q21	AR	Deletion	Naxos disease	(Antoniades et al., 2006; Corrado et al., 2000)
			AD	Insertion	ARVC	(Asimaki et al., 2007)

AR = Autosomal Recessive
AD = Autosomal Dominant

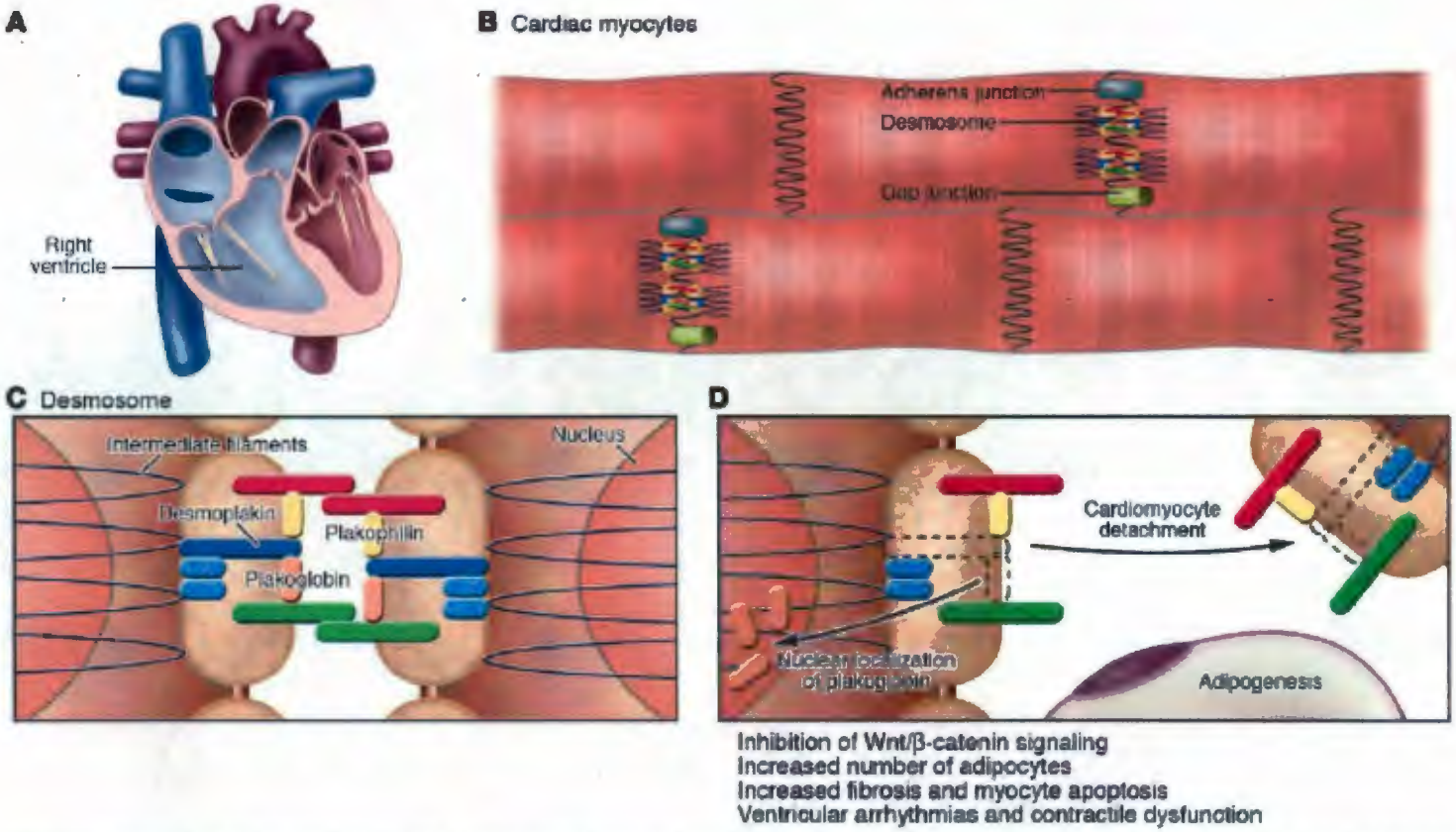


Figure 1: Cardiac-specific restriction of the desmosomal protein desmoplakin causes nuclear localization of plakoglobin and reduced Wnt/β-catenin signaling, recapitulating human ARVC. **(A)** ARVC predominantly affects the right ventricle of the heart. **(B)** The intercalated discs of cardiac myocytes are characterized by gap junctions, adherens junctions, and desmosomes. **(C)** Cell-cell adhesion is largely dependent on interaction of intracellular components of the desmosomal plaque such as desmoplakin and plakoglobin. **(D)** In this issue of the *JCI*, Garcia-Gras et al. (16) report that heterozygous cardiac desmoplakin-deficient mice show nuclear localization of plakoglobin and reduced Wnt/β-catenin signaling. This causes increased expression of adipogenic and fibrogenic genes in vitro, abnormal cardiac adipose tissue and fibrosis in vivo, and ventricular arrhythmias similar to human ARVC. Interactions between signaling defects and mechanical stresses are likely to be involved in the genesis of the final phenotype.

Reprinted with permission from: (MacRae (2), Birchmeier, & Thierfelder, 2006)

1.2 GENE MAPPING AND IDENTIFICATION

1.2.1 History

In 1913, Thomas Hunt Morgan and his students made the first genetic/linkage map of the fruit fly. This early map gave only the relative positions of genes but not their physical location nor base pair distance between them. Another limitation was that they were based on the inheritance of traits, not genes, which relied on phenotypic characteristics (Morgan, Sturtevant, Muller, & Bridges, 1922; 1922). In 1945, Salvador E. Luria and Alfred D. Hershey independently demonstrated that viruses can mutate. A year later, Hershey discovered that when different strains of viruses infect a bacterium, the viruses could exchange genetic material (Hershey, 1952; Luria, Fraser, Adams, & Burrous, 1958). In the 1970s, Stanley Cohen and Herbert Boyer initiated the development of recombinant DNA technology which allowed small pieces of DNA to be inserted into bacteria. These bacteria could be cloned and grown in large quantities (Cohen, Chang, Boyer, & Helling, 1973). This new technology was the springboard needed for much advancement in molecular genetics and biochemistry. Over the next two decades, new techniques for mapping and rapidly sequencing genes have fueled the development of genetic engineering.

Despite intense worldwide research of human chromosomes and the genes located on them, there may be more genes to identify. The function of many of the genes already identified and their expression control mechanisms has yet to be discovered. The human

nuclear genome consists of 22 autosomes and 2 sex chromosomes, X and Y. In 2001, the International Human Genome Sequencing Consortium reported a draft sequence of the euchromatic portion of the human genome. They estimated that 30,000 human protein-coding genes exist (Lander et al., 2001). In 2004, the scientific description of the finished human genome sequence was published. It was estimated that there were three billion nucleotide base pairs containing 20,000-25,000 human protein-coding genes (International Human Genome Sequencing Consortium, 2004). A recent version of the human gene catalogue estimated that the human genome consists of over 3.25 million base pairs containing 21,541 protein-coding genes with 275,708 gene exons (Ensembl 36, Oct 2005). The remaining non-exonic genomic DNA is made up of RNA genes, introns, regulatory elements, and so-called junk DNA, for which no function has yet been identified. Interestingly, the presence of many additional conserved sequences suggest that they have undetermined functional importance.

1.2.2 Genetic/Linkage Maps

Genetic maps, also called linkage maps, are critical for identifying chromosomal locations and associations of two or more genes or genetic markers. They define the position of genes on a chromosome relative to each other in terms of recombination frequency and aid in identifying mutations responsible for heritable phenotypes. The genetic distance separating the two loci is directly proportional to the recombination frequency. The accuracy of this estimation increases as the physical distance, the gap between two loci measured by the number of base pairs separating them, contained within

the loci decreases. Increasing the number of markers on a map increases the probability that a gene of interest located on a specific chromosome will be closely linked to a marker.

Markers or genes having multiple alleles are termed polymorphic and can be used in genetic studies to investigate segregation of traits in families. These markers are found in nuclear and organellar DNA and consist of repeating units with a length of 1 to 4 base pairs (Taylor, Day, & Human Genome Organisation, 2005; Turnpenny, Ellard, & Emery, 2007; Watson & Watson, 2007). Polymorphic markers are considered highly polymorphic if they contain 6 or more alleles and display 75% or more heterozygosity (Luty et al., 1990). Such markers were instrumental in the construction of detailed genetic maps (Andersson et al., 1994). Microsatellite markers are used in a wide range of applications, such as: linkage and association studies looking for disease loci; population studies; determination of kinship and gene dosage; and, forensic identification. Thousands are dispersed throughout the human genome, making them invaluable for acquiring genetic information (Matise et al., 2007; O'Brien et al., 1999).

The underlying principle in genetic mapping is the coinheritance of marker alleles and gene mutations which can be followed within families. Two loci are linked if their alleles consistently cosegregate in successive generations. If genes are linked, then the traits they control will be most likely inherited together. These cosegregating linked genes are commonly referred to as linkage groups (Botstein, White, Skolnick, & Davis, 1980;

Watson & Watson, 2007).

Recombination plays an important role in creating genetic diversity within species, and inferring past recombination events is central to exploring many problems in genetics. When a crossover occurs between two loci, the alleles at the loci are considered to have undergone recombination. Recombination is the process whereby there is a formation of new combinations of alleles in offspring as a result of the exchange of DNA sequences between sister chromatids. This process occurs during meiosis of the cell cycle. A chromatid strand may break and join a different chromatid - commonly its homologous chromosome. Thus, meiotic recombination will cause cosegregating genes to be separated if a recombination event occurs between them. If a recombination event occurs, the chromosome and genes contained in the offspring may differ from that of its parents (Haines & Pericak-Vance, 2006).

Several factors can influence the number of recombination events on a specific chromosome. One such factor is gender. In humans, genome-wide genetic maps of females are always significantly longer than male maps because recombination rates in female meioses are approximately 1.6 to 1.7 times greater than in male meioses (Lynn, Ashley, & Hassold, 2004). Another influencing factor is that of chromosomal location. For example, a higher rate of recombination is seen in telomeric regions of males when compared to female subjects. Also, regardless of gender, there are other genomic regions, such as the centromere, which have fewer recombination events than the genome average

(Choo, 1998). Genomic regions with relatively low recombination frequency are referred to as *coldspots*, while regions exhibiting much higher rates of recombination than average are known as *hotspots*. It has been estimated that the human genome contains over 25,000 hotspots (Myers, Bottolo, Freeman, McVean, & Donnelly, 2005). Analyses of individual chromosome regions and genome-wide studies suggest the presence of interindividual variation (Laurie & Hulten, 1985; Lynn et al., 2002; Yu et al., 1996). Surprisingly, age does not appear to change the rate of recombination in either human females or males (Lynn et al., 2000; Shi et al., 2002).

1.2.2.1 Genotype and Haplotype Analysis

A number of concepts, such as genotype and phenotype, are important in gene identification. A *genotype* is defined by the pair of alleles at a given locus on homologous chromosomes. A *phenotype* is the observable expression of a gene and the organism's observed characteristics. An organism's phenotype is influenced both by its genotype and by its environment. Evolution can change a population's genotypes from one generation to the next, which, in turn, can change the population's phenotypic characteristics. The pattern of inheritance reveals how genotypes are associated with the phenotypes. The two most common patterns of inheritance with autosomal chromosomes are dominant and recessive. Autosomal dominant inheritance occurs when only one copy of a mutant allele causes an expression of the disease. Autosomal recessive inheritance requires two copies of a mutant allele for expression.

In recent years, fine-mapping has proved a valuable method for ancestral haplotype reconstruction (MacLean et al., 2000; Morris, Whittaker, & Balding, 2000; Morris, Whittaker, & Balding, 2002). A haplotype is a sequence of DNA (or alleles) on a chromosome that is transmitted together to the offspring because they are physically linked to each other. Gene mapping can be conducted by searching for a conserved haplotype shared by multiple patients (Houwen et al., 1994). The probability that a haplotype is identical-by-descent increases rapidly as the number of shared alleles increases (Merriman et al., 1998; Nolte & Te Meerman, 2002). In haplotype sharing analysis, we identify genomic regions that have shared haplotypes in individuals with similar phenotypes. It is often measured as the shared length, or the number of shared allele intervals in haplotypes being compared. It has been shown that the shared haplotypes in unrelated cases can be used to localize disease genes (te Meerman & Van der Meulen, 1997; Van der Meulen & te Meerman, 1997). We used this approach to estimate the disease gene location and the ancestral haplotype.

Haplotypes can increase the information for linkage with a set of markers, which otherwise may provide weak information if these markers were individually assessed. By evaluating all linked markers simultaneously, the accuracy and informativeness of the haplotype increases (Du, Woodward, & Denise, 1998). Also, typing additional markers within a haplotype increases the informativeness of the haplotype and the probability that the haplotype does not occur in the general population. This provides support for the theory of an ancestral haplotype.

1.2.2.2 Fine-mapping

The initial linkage of a diseased gene locus to a genomic region does not identify the disease-causing gene. Fine-mapping involves the identification of markers that are very tightly linked to a diseased gene. When parental genotypes are available, a map density of one microsatellite marker per 2 cM is recommended. However, an increase in marker density should be applied when parental genotypes are not available. A sufficient marker density will aide in acquiring the maximum amount of inheritance information from any panel of markers investigated (Evans & Cardon, 2004). Hence, by incorporating additional markers into the marker set, an increase in identified meiotic recombination events will reduce the locus surrounding the diseased gene, thus, increasing the resolution. Fine-mapping a gene is usually an essential step in map-based gene isolation, whereby the researcher can acquire the gene for further meticulous study (Andersson et al., 1996; McPeek, 2000; Slatkin, 2008).

1.2.3 Gene Identification

Two general approaches may be used when attempting to identify a disease-causing gene associated with a Mendelian disease. The first is *functional screening* (candidate gene approach), by which genes are selected according to their known function and pattern of expression, regardless of their genetic map location (Kwon & Goate, 2000). The other approach is *positional cloning*, a method used to identify a causative gene when its approximate chromosomal location is known from linkage analysis in the absence of

information regarding its protein product (Watson & Watson, 2007). This method is used when linkage mapping has identified a diseased-linked region. Ideally, the candidate region will be flanked between two markers that are tightly linked to the unknown disease-causing gene. In positional cloning, all known genes within the identified disease-associated locus are identified with the use of public databases (e.g. NCBI and UCSC) and selected for mutation detection. Published Open Reading Frames (ORFs) provide researchers with a normal reference sequence that encodes a gene product. Primers, or probes, are designed by selecting sequences from the ORFs that are complementary to the genomic region to be sequenced. Next, genomic regions of affected individuals carrying the mutant DNA are sequenced and compared to the normal reference sequence. This mutational analysis identifies potential disease-causing candidate genes when the mutant DNA from affected individuals differs from the normal reference sequence. For this study, the positional cloning approach was utilized.

1.2.3.1 Identifying Pathogenic Variants

Once a candidate region is identified through linkage analysis and narrowed by fine-mapping, the genes residing within this region are deemed positional candidate genes. In recent years, DNA segments of genes within the candidate region have been isolated using cloning techniques; and, computer programs (e.g. Mutation Surveyor) have been used to analyze sequencing data to screen these positional candidate genes for causative mutations. All gene sequences produced must be compared to readily available wild type DNA sequences in order to identify the mutation that causes the disease. Functional

candidate genes which reside within the candidate region and are known to be expressed in the affected tissue or functional pathways are prioritized for mutation screening (Watson & Watson, 2007). Prioritizing positional candidate genes according to their function associated with the disease phenotype can reduce the time required to identify the mutation.

Finding a single disease-causing gene by comparing affected with unaffected individuals is a straightforward process compared to identifying multiple genes causing a single disease. Researchers can theorize that a disease-causing gene is near a specific marker allele when that allele is found more often in affected individuals than in unaffected individuals. As mentioned above, genes in this region are considered positional candidate genes. A genomic difference between affected and unaffected individuals has proved beneficial in identifying disease-causing mutations, especially in rare inherited diseases such as cystic fibrosis (Riordan et al., 1989) and myotonic dystrophy (Salvatori et al., 2005).

An important component in the research process is that of population controls - individuals who live in the same geographical region as the patients being investigated and, usually, are ethnically matched. They do not have the disease and/or are randomly selected from the population. The use of population controls is beneficial when comparing patient characteristics (e.g. a genetic variant) to the unaffected population,

since a difference in the frequency of the variant between the two groups indicates that there may be an association of the variant with the trait.

1.3 GENE MAPPING IN THE NEWFOUNDLAND POPULATION

In the present era of genomic revolution, isolated populations are important resources for biomedical research due to their secluded nature, low rates of migration, and reduced environmental and genetic diversity (Shifman & Darvasi, 2001). The Newfoundland population is one such isolated population, whose genetic homogeneity has proven advantageous for the identification of gene mutations causing Mendelian disorders.

1.3.1 Founding of the Newfoundland Population

Newfoundland's historical patterns have created a genetically isolated homogeneous population – both within the population as a whole and within community sub-isolates (Rahman et al., 2003). The favorable characteristics (e.g. isolation, ethnic homogeneity, and consanguineous marriages) of the Newfoundland population (Mannion et al., 1977) provide researchers with a resource for investigating genetic diseases and their underlying causes. The founding population of Newfoundland originated from Great Britain and Ireland. After taking possession of Newfoundland in 1583, Britain soon began establishing year-round settlements, with peak immigration to Newfoundland occurring in the mid-1700s. The founding population consisted, mainly, of Protestant settlers from the south-west of England and Roman Catholic settlers from the south of Ireland (Mannion et al., 1977).

Over the following centuries, the population of Newfoundland had risen to 568,474 (Statistics Canada 1991) with a subsequent declined to 505,469 in 2006 (Statistics Canada 2006). Because the original founding communities were isolated, they experienced little immigration (Bear et al., 1987). Thus, they are ethnically homogeneous and their inhabitants share a similar lifestyle and environment. This isolation sustained large families with little dispersion and promoted elevated consanguineous marriage rates (Bear et al., 1988).

Genetically isolated populations have proved to be extremely important for mapping and identifying genes involved in Mendelian conditions. The degree of variation in homogeneous populations is relatively low. This is due to greater genetic homogeneity and fewer environmental factors affecting it (Heutink & Oostra, 2002; Peltonen, Palotie, & Lange, 2000). This lower degree of variation allows the experimental design to incorporate smaller sample sizes when samples are limited. This proves to be beneficial when the study is investigating a weak genetic contribution or a rare genetic condition (Heutink & Oostra, 2002). The Newfoundland population is genetically isolated, relatively homogeneous, and exhibits a low degree of variation. Isolated populations, such as the Newfoundland population, are advantageous to mapping genes for diseases that are frequent in that population.

1.3.2 Factors Affecting Genetic Disorders in the Isolated Population

In approaching a specific population for genetic research, researchers should be aware of

the various factors that may affect the degree of homogeneity, such as: the founder effect; bottleneck effect; genetic drift; timing of immigration and growth spurts; and, recombination events.

The founder effect occurs when a small genetically isolated population experiences higher incidence rates of genetic diseases after it migrates from a larger population. While some of the diseases present in the parental population are found in the smaller founding population, not all diseases will manifest (Mayr, 1963). The reduction in genetic variation in the small founding population, in comparison to its parental population, can result in a population that is both genetically and phenotypically different from its parental population (Ridley, 2004).

The founder effect is substantiated by disease alleles showing inheritance *identical-by-descent* (Botstein & Risch, 2003; Lander & Botstein, 1987) which occurs when two alleles at a single locus, in individuals whose relation is unknown, are identical copies of the same alleles from a common ancestor, without any intermediate mutation. In genetically isolated populations, a common ancestor of the founding population can pass alleles down distinct lineages (Grant, Manduchi, Cheung, & Ewens, 1999). One example of the prevalence of a disease being influenced by the founder effect was found in research on myotonic dystrophy, in the northern Transvaal, South Africa. The minimum prevalence of this disease in the South African Caucasoid Afrikaans-speaking families is 14.3/100,000 (Lotz & van der Meyden, 1985). This is approximately 3 times the

prevalence in more admixed Caucasian populations of European ancestry (2.4-5.5/100,000) (Moxley, Meola, Udd, & Ricker, 2002).

Another factor affecting the degree of homogeneity in the isolated population is known as the bottleneck effect, which results when there is a significant reduction in the size of a population, followed by population growth (Keller et al., 2001). A low rate of immigration during a population's growth period leads to an increase in consanguinity, as experienced by the Newfoundland population.

Genetic drift results from the random change in allele frequency within a population as genes are passed from one generation to the next. It is dependant on allele frequency and population size. For example, if human parents have a small number of offspring, than not all of the parent's alleles will be passed on due to chance assortment of chromosomes at meiosis. The effect on allele frequency will be minimal in each generation of a large population because the random nature of the process will tend to average out. However, in a small population the effect of this random allelic fluctuation does not average out, making the population less likely to evolve and adapt to new selective pressures. Therefore, the allele frequency may change rapidly and over time may become homogeneous if the population size doesn't change (Peltonen et al., 2000). This is true for both recessive and dominant alleles, with the exception of alleles associated with early-onset disorders which can reduce the fitness of carriers to reproduce. Early-onset disease alleles will be lost from an isolated population much faster than those lost in the

general no isolated population due only to natural selection (Botstein & Risch, 2003). Genetic drift is unlike natural selection, in which an organism's genotypic characteristics increase the possibility of the organism adjusting to environmental changes which, in turn, increases survival and reproduction. Natural selection cannot increase the frequency of a characteristic once it has disappeared from the population - it can only utilize the genetic variation already present. A population may have difficulty adapting to new selection pressures (e.g. climate change) when the genetic pool is reduced in that population. This is due to the genetic variation, which the selection would call upon, having already been drifted out of the population.

The degree of homogeneity resulting from the bottleneck effect and genetic drift will be greater if immigration is kept to a minimum prior to population growth (Peltonen et al., 2000). Young population isolates (<20 generations), including the Newfoundland population (Rahman et al., 2003), the Hutterites (Hostetler, 1985) and the Central Valley of Costa Rica population (Morera & Barrantes, 2004) have shown an exponential growth pattern in recent generations. Out migration during the past generation has resulted in a decline of the Newfoundland population from 568,474 in 1991 to 505,469 in 2006 (Statistics Canada 1991 and 2006). Genetically isolated populations allow us to investigate the influence of the bottleneck effect and high levels of genetic homogeneity when conducting genetic research of Mendelian disorders (Heutink & Oostra, 2002).

A factor reducing homogeneity is recombination events between heterogeneous chromosomal segments, which tend to increase the variation of a population over time (Zavattari et al., 2000). As previously mentioned, recombination results in an exchange of segments between paired chromosomes undergoing separation. During meiotic recombination, the chromosomes from a person are shuffled to create a collage of genetic material. The unique set of chromosomes created can be inherited by a person's offspring, assuming no additional recombination event occurs at this stage. This is the source of variation in populations. For example, older isolated populations (>100 generation), such as the Sámi in northern Scandinavia, have been described as having more genetic variation than the younger isolated populations of Finland, Iceland, Sardinia, Japan and Newfoundland, due to the longer period of time during which recombination events have occurred (Heutink & Oostra, 2002; Kruglyak, 1997).

1.3.3 Benefits of Identifying Founder Mutations in Isolated Populations

Some cases of genetic disorders in a genetically isolated population are the result of one or a few founder mutations. This is efficient for diagnostic testing because few mutations will have to be screened in patients from the isolated population. For example, founder mutations in the APC and MSH2 genes have been identified as the cause of colorectal cancer (CRC) in the Newfoundland population. A total of four founder mutations have been found in 32 high-risk families – two in APC and two in MSH2 (Green et al., 2007). The incidence of CRC in the Newfoundland population is 27% higher then the Canadian average (National Cancer Institute of Canada (2003) Canadian Cancer Statistics 2003.

Toronto, Canada). As a result of this discovery, the diagnosis of patients presenting with CRC can start with the screening of these mutations prior to screening the entire APC and MSH2 genes.

1.4 PROPOSAL OF RESEARCH

1.4.1 Rationale

In this research, the underlying genetic basis of ARVC linked to the ARVD5 locus on chromosome 3p25, in the Newfoundland population, was determined. The main purpose of this project was to narrow the ARVD5 critical region and to identify the disease-causing gene. This would provide insight into ARVC in general; new clues to the biochemical pathways that are involved in the etiology of ARVC; an opportunity for early and more accurate molecular diagnosis; and, possibly a more effective treatment for affected and at-risk individuals.

1.4.2 Objectives

1. To reduce the candidate ARVD5 locus by fine-mapping critical recombination events
2. To identify genes within the refined candidate region by use of genome databases
3. To identify sequence variants of genes within the candidate region by DNA sequencing
4. To verify the pathogenicity of variants by comparison to population controls and by analysis of intra-familial segregation

1.5 CONTRIBUTION TO RESEARCH

Clinical ascertainment, counseling and treatment of patients were carried out by a full clinical team, not including me. It included, and is not exclusive to: cardiologist Dr. Connors; clinical epidemiologist Dr. Pat Parfrey; and, genetic counselors Kathy Hodgkinson and Sara MacKay.

The genotyping team consisted of Dante Galutira, Ingrid Pardoe and me. Dante taught me the polymerase chain reaction assays and helped carry out a large portion of this procedure during the genotyping phase of my research. Ingrid's many lab duties included DNA bank management for my project. Both Dante and Ingrid worked with me during the first half of the genotype data analysis process. I carried out the remaining polymerase chain reaction assays, genotype data analysis, importing existing Cyrillic pedigrees and genotype data into Progeny software, and all manual haplotype construction.

The molecular sequencing team consisted of Nancy Merner, Dr. Annika Haywood, and me. Nancy performed as our team leader. Nancy, Annika, and I sequenced and analyzed 20 positional genes. Of the 20 candidate genes investigated, I had fully or partially carried out primer work-up on 10 genes, sequencing with ARVC test samples on 9 genes, initial data analysis on 2 genes, and verification of data on 6 genes. Nancy and Annika used population controls to determine the variant frequency for each of the variants segregating on the ARVD5-associated haplotype. Nancy carried out the segregation

analysis of rare variants in three affected families. Annika and Fahad Chowdhury, a summer student, carried out bioinformatics analysis on the identified gene and mutation.

A manuscript containing some of my thesis research has been published. Nancy D. Merner, Kathy A. Hodgkinson, Annika F.M. Haywood, Sean Connors, Vanessa M. French, Jo¨rg-Detlef Drenckhahn, Christine Kupprion, Kalina Ramadanova, Ludwig Thierfelder, William McKenna, Barry Gallagher, Lynn Morris-Larkin, Anne S. Bassett, Patrick S. Parfrey, and Terry-Lynn Young. (2008) Arrhythmogenic Right Ventricular Cardiomyopathy Type 5 Is a Fully Penetrant, Lethal Arrhythmic Disorder Caused by a Missense Mutation in the *TMEM43* Gene. The American Journal of Human Genetics 82(4), 809-12, April 2008.

CHAPTER 2: SUBJECTS AND METHODS

This section outlines the process that was followed to identify the gene and mutation causing ARVC at ARVD5. These steps are summarized in the flowchart of Figure 2.

2.1 STUDY POPULATION

Of the 152 families referred to the provincial genetics program due to a family history of cardiomyopathy and/or sudden cardiac death (SCD), 15 families had individual(s) diagnosed with SCD ≤ 50 years of age and were investigated in this study. Family members, or their next of kin, consented to participate in the research, which was approved by the Human Investigations Committee of the Eastern Health Corporation of St. John's, Newfoundland, Canada (study number 00-176).

One of the 15 families, family 64, was the original family in which linkage to chromosome 3p23 was defined (Ahmad et al., 1998). Subsequently, a collaborative effort with Dr. Ludwig Thierfelder's research team at the Max-Delbrück Centrum für Molekulare Medizin, Berlin, confirmed that the remaining 14 families were also linked to the same region (unpublished data) - 19, 69, 76, 185, 273, 453, 581, 840, 853, 864, 932, 964, 977 and R1139.

2.2 CLINICAL INVESTIGATION

Approximately 250 members of the 15 NL 3p25-linked families were assessed. Each family was included in our ARVC study if their history included individual(s) diagnosed with SCD ≤ 50 years of age.

2.2.1 Clinical Criteria

Clinical diagnosis was used to define the ARVC affection status of family members in pedigrees – primary, secondary, and unaffected/unknown. This classification of affection status allowed us to quickly review the severity of a family member's known ARVC clinical signs within the pedigree. Not all family members underwent clinical assessment.

Primary affection status was applied when one or more of the following was true:

1. SCD ≤ 50 years of age.
2. Obligate Carrier (preferably parent of an affected child, but in some circumstances, an individual with an affected child and sibling).
3. Ventricular tachycardia (VT), with or without cardioversion ≤ 50 years of age.
4. Dilated cardiomyopathy (DCM) and/or congestive heart failure (CHF), with or without heart transplant ≤ 50 years of age.
5. Replacement of myocardium, on autopsy, with fat and fibrous tissue.

Secondary affection status was applied when one or more of the following was true:

1. SCD otherwise unexplained > 50 years of age.
2. Non-sustained VT on Holter or telemetry.
3. Ectopy with premature ventricular contractions (PVC) > 1000 on Holter or telemetry.

Unaffected/unknown status was defined as one of the following criteria:

1. Clinical assessment not meeting primary or secondary status criteria.
2. No clinical assessment data available.

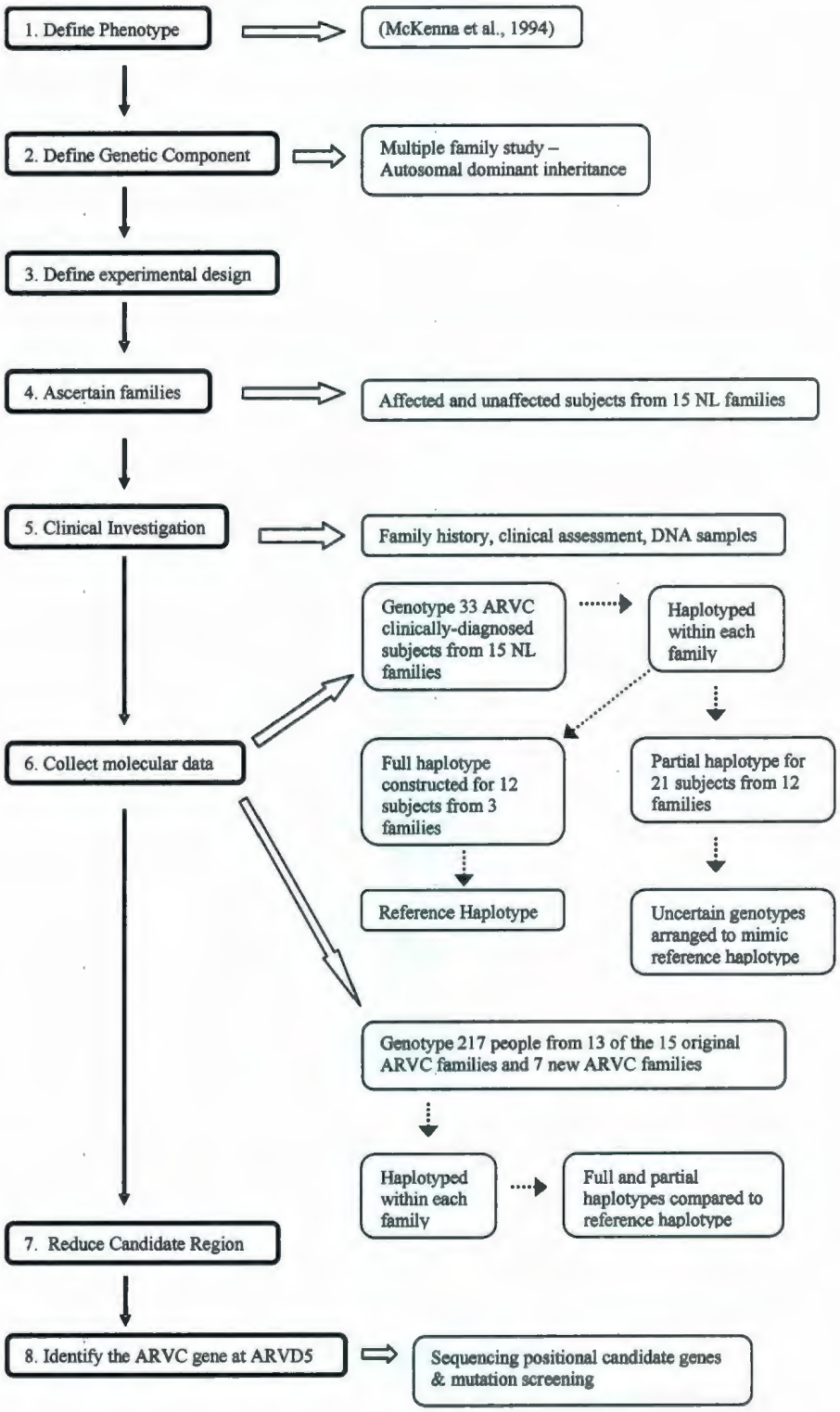


Figure 2: Flowchart of gene discovery process using 15 ARVC families linked to 3p25 (ARVD5).

2.3 MOLECULAR METHODS

2.3.1 Genotyping ARVC families with polymorphic markers on 3p25

Genomic DNA from approximately 250 members of the 15 Newfoundland families, including family 64, was used. To narrow the previously described ARVD5 9.93 Mb critical region, 17 genetic markers were used - *D3S3610, D3S2403, D3S1516, 3S3608, D3S2385, D3S3602, D3S1585, D3S1554, D3S3595, D3S3613, D3S3473, D3S2338, D3S4547, D3S3510, D3S1293, D3S3038,* and *D3S3659.* The average physical distance amongst the 17 markers was 0.64 Mb, ranging from 0.015 – 2.8 Mb. (Table 3). Eight of the markers were originally used to map the ARVD5 region (Ahmad et al., 1998); and, the remaining 9 were added to increase resolution.

Fluorescently-labeled forward markers were used (supplied by ABI Applied Biosystems). The fluorescent dyes were 6-FAM (blue), PET (red), HEX (yellow), and TET (green). Reverse markers were unlabelled (supplied by IDT Integrated DNA Technologies). Genotyping was performed on an ABI PRISM model 3100/3130 analyzer (Applied Biosystems). Fluorescent dyes, of differing colors and different fragment sizes, permitted analysis of multiple genetic markers within a single electrophoretic run.

Table 3: Markers on chromosome 3p25 used for fine-mapping the ARVD5 locus
Information gathered from the UCSC Genomic Bioinformatics website (http://genome.ucsc.edu/cgi-bin/hgGateway). Assembly: March 2006.

Marker	Genomic Start	Genomic End	Forward Sequence	Reverse Sequence	Variation Type	Size Range (bp)	Maximum Heterogeneity	Marker Type
D3S3610	12,980,656	12,980,996	CCAGATTCTCTA AGGCCATGC	TGTGGAATATCC GCCCAG	dinuc	211-263	0.69	Microsatellite
D3S2403	13,147,397	13,147,709	ACAGATTGAGAC CATGTGTCA	CACACTCAAAAT ACATGAAGGC	tetra	247-283		Microsatellite
D3S1516	13,628,628	13,629,103	TAGTCCCACCTA CTCGGGA	AGGACCCTGTTG ACAGGGT	tetra	331-367	0.8571	Microsatellite
D3S3608	13,679,236	13,679,541	ACGCCATCTNCA GGCT	TGGGTGCAGCAC ACCA	dinuc	168-196	0.75	Microsatellite
D3S2385	13,853,945	13,854,287	GCTGTATTCGGG AGCATCTA	CCACCATGAAAG AATGGCTA	tetra	142-154		Microsatellite
D3S3602	13,900,968	13,901,215	AAAATCCTAACC CAAAATGT	ATCAGAAAATAA CAGAGGGC	dinuc	114-132	0.5106	Microsatellite
D3S1585	13,916,682	13,916,861	TGCACGAGCCAG AAGT	TTGGACTGCTGA GGGG	dinuc	126-144	0.59	Microsatellite
D3S1554	14,342,778	14,343,106	ATTCATCTTGTT ACTGTTCATTTG T	GGGCAAACCCAA AGACT	dinuc	133-141	0.55	Microsatellite
D3S3595	14,617,332	14,617,642	AAATTCAAATGT GGGGCACC	CTTGCTCTTGACA CCCCCTG	dinuc	269-273	0.2947	Microsatellite
D3S3613	15,336,926	15,337,204	CATCTATGTGGC AATCGG	CAGCATTTGTTGT AGGGACT	dinuc	172-208	0.8522	Microsatellite
D3S3473	16,497,252	16,497,602	TCAAGGCTGCTT TTATGAATCC	TACTCAGTCTTGT ATGAAAATGGGC	dinuc	217-225	0.7028	Microsatellite
D3S2338	16,824,399	16,824,646	GAAGCCAGCAGT TTCTC	CTGTATTGTTTTC CAGGATAAG	dinuc	179-197	0.8883	Microsatellite
D3S4547	17,927,223	17,927,699	AACTGGTTTTGA CTGGACCA	AGAAGCAAGAGG AAGAAGGC	tetra	221-248		Microsatellite
D3S3510	19,073,530	19,073,906	GGACTATCACAA AGTATGGCA	GGGAAATGAACA AATCACAT	dinuc	279-289	0.7243	Microsatellite
D3S1293	21,902,053	21,902,384	ACTCACAGAGCC TTCACA	CATGGAAATAGA ACAGGGT	dinuc	116-144	0.7979	Microsatellite

D3S3038	21,924,404	21,924,719	CATCTTTCTTTTC CTGTTCCC	GATACCATATTCA ACATGAAGAGG	tetra	187-219		Microsatellite
D3S3659	22,913,709	22,914,093	TACTTCATGCAA AGAATCTACCA	CGAAGTTAATCA GGGAAGCA	dinuc	248-260	0.6517	Microsatellite

2.3.1.1 Amplification of DNA by Polymerase Chain Reaction (PCR)

Each standard PCR reaction included: 1X PCR buffer; 1.5 mM $MgCl_2$; 0.2mM of each dNTP; 20-100 ng of DNA template; 0.5 U of *Taq* polymerase; and, 0.4 μmol of each primer (Invitrogen, USA). When necessary, 0.0878 g/mL of betaine and autoclaved distilled water were added to bring the mixture to a total volume of 25 μL. Betaine can be very important in the PCR. It promotes amplification of high GC-containing target DNA, while also improving the amount of product and increasing the resistance of the polymerase to denaturation. It also permits amplification with lower quality DNA due to low levels of contaminants that often co-purify with DNA (Henke, Herdel, Jung, Schnorr, & Loening, 1997; Weissensteiner & Lanchbury, 1996).

Reaction mixtures were amplified by heating for 5 cycles to 94°C for 5 min 30 sec (denaturation step), 64°C for 30 sec (annealing step), and 72°C for 30 sec (extension/elongation step), followed by 30 cycle of 94°C for 30 sec, 54°C for 30 sec, and 72°C for 30 sec. Finally, tubes were heated at 72°C for 7 min (final elongation). This was performed on a 9700 GeneAmp PCR system (ABI Applied Biosystems).

To confirm amplification, 4-5 μL of PCR product plus 1 μL of loading dye, along with a 100bp size ladder and a negative control, were run on a 1% agarose gel containing ethidium bromide. The bands were visualized using an ultraviolet transilluminator. When amplification was not achieved with standard PCR conditions, various PCR reactions and conditions were modified to achieve amplification (see Appendix A).

For the preparation of a 1% agarose 40 mL gel, 0.4 g agarose and 40 mL 1X TBE were added to a 250 mL beaker, mixed and heated in a microwave for 1 minute. 2 μL of ethidium bromide (10 mg/mL) was added and mixed into the hot gel before pouring into a gel tray for further polymerization. The same protocol was used for a 150 mL gel except 1.5 g agarose, 150 mL 1X TBE and 7.5 μL ethidium bromide were added to a 500 mL beaker. 1 μL of loading dye; plus, 4-5 μL of PCR product was loaded on a 1% agarose gel. Fifty microlitres of 100bp DNA ladder was mixed with 950 μL of distilled water to make a working solution. 3 μL of this diluted ladder mixed with 1 μL of loading dye was loaded in each lane of the agarose gel.

2.3.1.2 Preparing Genotype Fragments for Analysis

A varying amount, 0.5 – 0.75 μL, of PCR product was added to 0.5 μL internal size standard (GeneScan -500 LIZ Size Standard) and brought up to a final volume of 10 μL with deionized formamide. Where possible, the PCR products for each DNA template were pooled. This is achieved by co-loading multiple PCR reactions, of different product sizes or labeled with different dyes if product sizes overlap, in the same capillary injection. The reactions were denatured and analyzed by capillary electrophoresis with the use of the ABI PRISM model 3100/3130 analyzer (ABI Applied Biosystems). The data was inspected manually and analyzed by Sequencing Analysis v5.2 and GeneMapper v3.7 software (Applied Biosystems).

2.3.2 Haplotype Construction

All pedigrees utilized in the research had been previously constructed using Cyrillic 2.1.3 (FamilyGenetix Ltd). Cyrillic pedigrees and associated clinical information were then imported into the genetic software Progeny v6.9.03 (Progeny Software, LLC). Due to the large volume of data, genotype data for multiple samples and genetic markers were assembled in MS Office Excel and imported into the Progeny program. To verify no data entry errors occurred, the original data was compared to the imported Progeny program data. In addition, proper genotype and haplotype segregation from parents to offspring was visually confirmed.

An individual's genotype was determined either by performing genotyping on their genomic DNA or by inferring from genotyped relatives. If there was at least one genotyped offspring for a genotyped parent, then that parent was considered genetically informative provided that the marker was heterozygous . However, an ungenotyped parent was only considered genetically informative if that parent had passed on both haplotypes to different offspring and the spouse was genotyped. A genotyped offspring was deemed genetically informative if there was at least one genotyped parent. A genetically informative individual was included in the genotyping analysis when their genomic DNA was available. This provided additional information during manual haplotyping procedure.

The marker order displayed on each pedigree reflects their physical order on the chromosome (http://genome.ucsc.edu/cgi-bin/hgGateway). We have assumed that the genotypes have undergone a Mendelian inheritance and that all pedigree relationships are accurate (i.e. correct parents, no adoption and correct DNA samples). Haplotypes were manually constructed by genotype comparisons, between and within families, while employing the minimum number of recombinants method. Missing alleles were inferred during the haplotyping process by their identical-by-descent copies. Haplotype segregation was examined manually within each family. Due to the large size of some pedigrees, individuals not useful for haplotype analysis were removed to produce core pedigrees for haplotyping (see Appendix B).

Given the large size of most pedigrees used, rules were sequentially applied to smaller subgroups (i.e. parent-offspring groups and nuclear families) within the pedigree until all genotyped individuals were manually haplotyped at all loci or until no individual could be further haplotyped at any locus. Several haplotyping rules were implemented throughout the procedure. First, when genotype data at a locus was missing in one parent but was known in the spouse and at least one offspring, then at least one allele in the missing genotype data could be inferred. Second, haplotypes were assigned to each offspring in accordance with the genotypes of their parents. Third, haplotypes were assigned to each parental locus conditional on the haplotypes of the offspring, and vice versa. Fourth, missing genotypes were assigned to each locus of the parent, conditional on the haplotypes of the offspring. In a similar manner, a genotyped individual may be

haplotyped if multiple siblings and other relatives are haplotyped. This procedure is applied repeatedly in a pedigree until no further changes can be made to either the genotypes or haplotypes in all individuals. Therefore, each individual is completely haplotyped at all loci, with the exception of uncertain loci (discussed below). In all cases, genotypes were arranged within a family such that the haplotypes produced had the minimum number of recombination events possible. When more than one haplotype produced the same recombination events, the haplotype was conservatively assigned the disease-associated haplotype at the ARVD5 locus. The remaining haplotype(s) was noted (data not presented). The pedigree subgroups were analyzed both from the eldest generation to the youngest generation and vice versa to reduce any bias from the direction of analysis (Guo, 1994; Wang, Xue, & Birdwell, 2006).

When genotypes at a locus could not be assigned clearly to a haplotype, the genotype was considered to be *uncertain*. For example, when a parent-offspring trio was identically heterogeneous at a locus and at least one parent and offspring had not been haplotyped, then that genotype was defined as uncertain in the unhaplotyped individuals. Also, if a parent's genotypes revealed more than one haplotype assignment at a locus that would result in an equal number of recombination events in the offspring, then the genotype was considered uncertain in the parent. Finally, if genotypes at a locus in an offspring could have been haplotyped more than one way to produce an equal number of recombination events, then the genotype was considered uncertain in the offspring.

2.3.2.1 Haplotype Analysis Within & Between Families

Initially, genotype data was generated on 33 clinically-affected individuals from 15 families using 17 genetic markers spanning the original 9.93 Mb ARVD5 critical region. Where possible, haplotypes were constructed for all individuals within their respective families. The original 9.93 Mb ARVD5 critical region shown in twelve fully haplotyped individuals was designated as disease-associated for all 17 genetic markers. There were two members of family 932, two members of family 1139, and eight members of family 64. This haplotype will be referred to as the ARVD5-reference haplotype. To illustrate the concept of haplotype sharing, genotypes in the remaining 21 individuals that could not be fully haplotyped were arranged to illustrate their consistency with the ARVD5-reference haplotype. Haplotypes and genotypes were visually compared across all families to assess what portion of the critical region was shared among the 33 clinically-affected individuals (see Results section, Table 4). Because not all carriers of the ARVD5 haplotype will experience cardiac symptoms, we used both obligate carriers and affected members to identify critical recombinations in the ARVD5 haplotype.

Other family members were later screened for the purpose of diagnosis, identification of recombination events, and comparing genotypes to the ARVD5-reference haplotype. These included additional clinically-affected patients, family members born at 50% *a priori* risk, patients with questionable clinical affection status, and some informative family members. There were a total of 217 additional people screened in 13 of the 15 original families (64, 69, 76, 185, 273, 453, 581, 840, 864, 932, 964, 977 and 1139) and 7

new families (34, 696, 920, 1143, 1144, 1173, and 1338). Ten markers spanning the critical region were utilized: *D3S3610, D3S2403, D3S1516, 3S3608, D3S2385, D3S3602, D3S1585, D3S1554, D3S3595,* and *D3S3613*. Where possible, haplotypes were manually constructed and recorded in Progeny. To illustrate the concept of haplotype sharing, genotypes in individuals that could not be haplotyped were arranged to illustrate their consistency with the ARVD5-reference haplotype.

2.4 SEQUENCING CANDIDATE GENES & MUTATION SCREENING

The strategy used in this study was to design primers covering all exons and untranslated flanking sequences of all 20 positional candidate genes, amplify the region by PCR, and verify appropriate amplification by gel electrophoresis. Each of the genes was bidirectionally sequenced with subset of affected and unaffected ARVC family members to identify polymorphisms unique to the affected patients. This sequencing panel consisted of 4 affected members and 3 unaffected controls - from families 64, 69, 453, and 840. Then segregation analysis was performed on patient-only polymorphisms within family 1139. Population control analysis identified rare variants (<1%). Finally, segregation analysis was performed in families 1139, 69 and 273 to identify a possible causative variant of ARVC at ARVD5.

The affected family members were good candidates for mutation screening due to their unambiguous clinical signs. They were beneficial for the confirmation of an identified sequence variation as a possible ARVC disease-related change at 3p25.

2.4.1 Automated DNA Sequencing of Candidate Genes

Using the University of California Santa Cruz (UCSC) Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway), a list of candidate genes for the ARVD5 locus was assembled for mutational analysis. For all positional candidate genes, PCR primer sequences were designed by PRIMER 3 (http://fokker.wi.mit.edu/primer3/input.htm) (Rozen & Skaletsky, 2000) to amplify

coding and non-coding exons, including at least 40 bases flanking each end of the exons. UCSC's BLAT program (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) was employed to quickly map each sequence, within the human genome, to verify its single unique position in the reduced ARVD5 critical region.

Each primer set was tested with the PCR conditions, described in section 2.4.1.1, using 2 random DNA samples (20-25 ng/μL each) and a blank, prior to analysis with ARVC test samples. The ARVC test sample DNA consisted of 4 ARVD5-linked (affected) and 3 unaffected (controls) family members (from 64, 69, 453, and 840). All samples were bidirectionally sequenced using primers and conditions described in Appendix C.

Sequencing fragments were evaluated using the ABI PRISM model 3730 analyzer (Applied Biosystems) at the CREAIT Centre, Memorial University of Newfoundland. The sequences were assembled and translated by Mutation Surveyor software version 3.00 (SoftGenetics, LLC). All sequences were compared to reference sequences (Lander et al., 2001).

2.4.2 Population Controls

Population controls were collected for the purpose of a colorectal cancer (CRC) study and supplied by Dr. Roger Green. They were age and sex matched with CRC patients and had no personal history of CRC.

CHAPTER 3: RESULTS

3.1 FINE-MAPPING OF THE ARVD5 REGION

Thirty-three clinically-affected members and obligate carriers of ARVC in 15 NL families were genotyped and, where possible, haplotypes were constructed within each family, based on minimum number of recombinants. If an individual's genotypes could not be haplotyped within their family, then the genotypes were arranged to illustrate their consistency with the ARVD5-reference haplotype constructed from families 64, 932 and 1139 (Appendix D2, D12 and D15).

A consistent 9.93 Mb haplotype (Table 4A, shown in yellow) that delineated in 8 affected members of family 64 and 2 affected members from each family 932 and 1139 was used as an ARVD5-reference haplotype for comparison to the other 21 clinically-affected patients' data (Table 4B and Appendix D1, D3-D11, D13 and D14). No single individual from family 64 had genotypes that could be fully haplotyped. However, a complete composite haplotype could be justified for family 64. All genotypes were haplotyped for individuals 64.0136, 64.A071 and 64.A073, except for marker D3S3602. However, genotypes for D3S3602 were haplotyped in individuals 64.0043 and 64.A142. Hence, by replacing the uncertain D3S3602 genotypes in individuals 64.0136, 64.A071 and 64.A073 with the haplotyped D3S3602 genotypes from individuals 64.0043 and 64.A142 a complete composite haplotype can be created for family 64. The haplotyped and uncertain genotypes for the remaining 3 family 64 individuals, 64.0003, 64.1034 and

64.0087, are arranged to match this composite haplotype. Hence, the composite haplotype for family 64 is incorporated into the reference haplotype (Table 4A).

Where possible, genotypes for the remaining 21 clinically-affected patients were haplotyped. The remaining genotypes were arranged to show their consistency with the ARVD5-reference haplotype. Comparison of 5 families (185, 581, 864, 964 and 977), with a total of 10 individuals, had genotypes consistent with the ARVD5-reference haplotype. Genotypes in the remaining 11 individuals (in families 19, 69, 76, 273, 453, 840 and 853) were only partially consistent with the ARVD5-reference haplotype (Table 4B). The various inconsistent genotypes eliminated 7.57 Mb of the original 9.93 Mb. Hence, all 33 clinically-affected members and obligate carriers, who were investigated, shared a telomeric 2.36 Mb portion of the previously described ARVD5 3p25 haplotype.

It is worth noting that individual 964.A021 was originally included in the analysis as an affected patient with only minimal clinical signs (see Appendix D13). She was later defined in the pedigree as clinically unaffected, due to her clinical signs not meeting primary or secondary affection status, as described previously.

The following section outlines the evidence of haplotype sharing found in the 12 partially haplotyped families and the strategy used to narrow the original 9.93 Mb.

Table 4: Haplotype sharing in the ARVD5 region of ARVC clinically-affected patients and obligate carriers.

Bolded genotypes were haplotyped within the family. Unbolded genotypes were uncertain.

A – Affected individual's genotypes that were haplotyped within their respective families. Yellow genotype blocks define the ARVD5 reference haplotype.

B – Affected individual's genotypes that were partially haplotyped. Genotypes that were not haplotyped were arranged to illustrate their consistency with the ARVD5-reference haplotype. Yellow genotype blocks are consistent with the ARVD5-reference haplotype. White genotype blocks are not consistent with the ARVD5-reference haplotype or did not reside on the disease-causing chromosome. Grey shaded markers were excluded from the critical region because its genotype, within one or more families, was not consistent with the ARVD5-reference haplotype. Unshaded markers define the reduced critical region where genotypes in all individuals were consistent with the ARVD5-reference haplotype.

A

	Families																							
	64																932				1139			
Marker	64.0043		64.A142		64.0003		64.A071		64.A073		64.0136		64.1034		64.0087		932.1000		932.0000		1139.0014		1139.1000	
D3S3610	**246**	**246**	**256**	**246**	242	246	**256**	**246**	**242**	**246**	**258**	**246**	**246**	**246**	**246**	**246**	**256**	**246**	**258**	**246**	nd	nd	nd	nd
D3S2403	279	251	279	251	283	251	**279**	**251**	**259**	**251**	**251**	**251**	279	251	**251**	**251**	**247**	**251**	**251**	**251**	**251**	**251**	**251**	**251**
D3S1516	**359**	**347**	**359**	**347**	343	347	**335**	**347**	**343**	**347**	**343**	**347**	**351**	**347**	**347**	**347**	**367**	**347**	**335**	**347**	**355**	**347**	**359**	**347**
D3S3608	**165**	**165**	**165**	**165**	175	165	**165**	**165**	**165**	**165**	**175**	**165**	**165**	**165**	**175**	**165**	**175**	**165**	**165**	**165**	**169**	**165**	**175**	**165**
D3S2385	**142**	**146**	**146**	**146**	138	146	**138**	**146**	**142**	**146**	**142**	**146**	142	146	142	146	**142**	**146**	**142**	**146**	**146**	**146**	**142**	**146**
D3S3602	**123**	**117**	**117**	**117**	123	117	123	117	123	117	123	117	123	117	123	117	**123**	**117**	**123**	**117**	**117**	**117**	**123**	**117**
D3S1585	**126**	**118**	**118**	**118**	126	118	**128**	**118**	**126**	**118**	**126**	**118**	128	118	120	118	**120**	**118**	**126**	**118**	**118**	**118**	**126**	**118**
D3S1554	**129**	**129**	**133**	**129**	**129**	**129**	**133**	**129**	**131**	**129**	**129**	**129**	**129**	**129**	135	129	**129**	**129**	**129**	**129**	**133**	**129**	**129**	**129**
D3S3595	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**	267	265	**265**	**265**	**265**	**265**	**265**	**265**	**265**	**265**
D3S3613	**197**	**193**	**193**	**193**	**201**	**193**	**197**	**193**	**179**	**193**	**199**	**193**	189	193	201	193	**201**	**193**	**191**	**193**	**201**	**193**	**191**	**193**
D3S3473	218	214	218	214	**214**	**214**	**214**	**214**	**212**	**214**	**218**	**214**	218	214	**214**	**214**	**220**	**214**	**218**	**214**	**218**	**214**	**218**	**214**
D3S2338	**185**	**185**	**185**	**185**	179	185	**175**	**185**	**187**	**185**	**177**	**185**	187	185	181	185	**175**	**185**	**183**	**185**	**181**	**185**	**179**	**185**
D3S4547	217	233	217	233	217	233	**225**	**233**	**217**	**233**	**229**	**233**	**233**	**233**	**233**	**233**	**221**	**233**	**225**	**233**	**221**	**233**	**225**	**233**
D3S3510	**283**	**285**	**285**	**285**	287	285	**283**	**285**	**283**	**285**	**287**	**285**	**285**	**285**	283	285	**281**	**285**	**283**	**285**	**283**	**285**	**283**	**285**
D3S1293	**8**	**8**	**12**	**8**	7	8	**5**	**8**	**2**	**8**	**12**	**8**	3	8	nd	nd	**7**	**8**	**2**	**8**	nd	nd	nd	nd
D3S3038	**4**	**4**	**5**	**4**	**4**	**4**	**4**	**4**	**1**	**4**	**4**	**4**	6	4	**4**	**4**	**5**	**4**	**0**	**4**	nd	nd	nd	nd
D3S3659	**2**	**4**	**6**	**4**	**4**	**4**	**5**	**4**	**3**	**4**	**4**	**4**	**4**	**4**	nd	nd	**4**	**4**	**6**	**4**	nd	nd	nd	nd

B

	Families																			
	19		69		76				185								273		453	
Marker	19.0000		69.A000		76.1011		76.0042		185.0000		185.1013		185.1011		185.1000		273.0000		273.A001		453.1012	
D3S3610	242	246	246	246	256	242	242	256	258	246	242	246	242	246	242	246	256	246	246	246	242	256
D3S2403	259	251	279	251	251	251	251	251	255	251	251	251	279	251	251	251	259	251	251	251	251	251
D3S1516	347	347	367	347	347	347	351	347	347	347	347	347	331	347	347	347	331	347	331	347	347	347
D3S3608	165	165	169	165	165	165	165	165	169	165	181	165	187	165	181	165	169	165	165	165	165	165
D3S2385	146	146	142	146	142	146	146	146	142	146	142	146	146	146	142	146	146	146	142	146	142	146
D3S3602	117	117	123	117	123	117	117	117	123	117	123	117	123	117	123	117	123	117	123	117	123	117
D3S1585	118	118	128	118	126	118	120	118	118	118	128	118	122	118	128	118	118	118	126	118	126	118
D3S1554	129	129	129	129	123	129	127	129	133	129	129	129	127	129	129	129	129	129	129	129	123	129
D3S3595	265	265	265	265	265	265	267	265	265	265	265	265	265	265	265	265	265	265	265	265	265	265
D3S3613	203	193	187	201	193	193	195	193	179	193	187	193	187	193	187	193	195	195	201	195	203	193
D3S3473	218	214	218	218	218	214	218	214	214	214	214	214	216	214	214	214	218	214	218	214	218	214
D3S2338	175	185	175	185	181	185	185	185	175	185	185	185	175	185	185	185	175	185	179	185	181	185
D3S4547	233	233	225	229	233	233	217	233	221	233	233	233	217	233	233	233	229	229	221	229	233	233
D3S3510	283	285	283	283	279	285	285	285	283	285	283	285	287	285	283	285	283	285	279	283	283	285
D3S1293	7	13	2	12	7	7	12	7	6	8	8	8	8	8	8	8	nd	nd	12	8	7	8
D3S3038	5	5	1	5	4	4	6	4	4	4	4	4	5	4	4	4	nd	nd	6	5	5	4
D3S3659	4	4	3	4	4	4	5	4	4	4	4	4	3	4	4	4	4	4	5	4	5	4

	Families																	
	581				840				853				864				964		977	
Marker	581.0001		581.0000		840.0000		840.0008		853.0011		853.0010		864.0000		864.1000		964.A021		977.0000	
D3S3610	256	246	256	246	256	246	256	246	256	246	256	246	256	246	242	246	246	246	256	246
D3S2403	251	251	251	251	251	251	251	251	251	251	251	251	251	251	251	251	251	251	279	251
D3S1516	331	347	331	347	nd	nd	351	355	355	347	355	347	351	347	355	347	331	347	nd	nd
D3S3608	165	165	165	165	169	165	165	165	165	165	165	165	183	165	193	165	169	165	165	165
D3S2385	146	146	146	146	146	146	146	146	146	146	146	146	142	146	146	146	142	146	146	146
D3S3602	121	117	121	117	111	117	117	117	117	117	117	117	125	117	123	117	123	117	123	117
D3S1585	118	118	118	118	134	118	118	118	118	118	118	118	118	118	122	118	128	118	126	118
D3S1554	129	129	129	129	129	129	129	129	129	129	129	129	129	129	129	129	129	129	129	129
D3S3595	265	265	265	265	265	265	267	265	267	265	267	265	267	265	265	265	265	265	nd	nd
D3S3613	199	193	199	193	179	193	197	193	197	193	197	193	197	193	187	193	199	193	201	193
D3S3473	218	214	218	214	214	214	214	214	214	214	214	214	218	214	214	214	218	214	nd	nd
D3S2338	181	185	181	185	181	191	181	187	187	185	187	185	181	185	181	185	175	185	nd	nd
D3S4547	233	233	233	233	229	229	217	229	217	233	217	233	229	233	225	233	229	233	229	233
D3S3510	283	285	283	285	283	283	279	283	279	285	279	285	283	285	285	285	285	285	nd	nd
D3S1293	2	8	2	8	2	7	12	7	12	12	12	12	7	8	8	8	8	8	nd	nd
D3S3038	0	4	0	4	0	4	nd	nd	5	5	5	5	6	4	6	4	7	4	nd	nd
D3S3659	4	4	4	4	5	5	5	5	3	5	3	5	5	4	5	4	4	4	nd	nd

3.1.1 Haplotype Sharing

Ten of 21 clinically-affected individuals had genotypes consistent with the ARVD5-reference haplotype, while the remaining 11 individuals were only partially consistent (Table 4B). The allele 246 for marker *D3S3610* identified in the ARVD5-reference haplotype was not observed in 2 individuals from family 76 and 1 individual from family 453. The alleles 347 and 185 for markers *D3S1516* and *D3S2338*, respectively, were not present in 2 individuals from family 840. One individual from family 69 and 2 individuals from family 273 did not carry the 193 ARVD5-reference allele for marker *D3S3613*. The same individual in family 69 did not carry the 214 ARVD5-reference allele for marker *D3S3473*. Alleles 185, 233 and 285 for markers *D3S2338*, *D3S4547*, and *D3S3510*, respectively, were not observed in 2 individuals from family 840; 5 individuals from families 69, 273, and 840; and 4 individuals from families 69, 273, and 840, respectively. The ARVD5-reference allele for *D3S1293*, 8, was not present in a total of 8 individuals from families 19, 69, 76, 840, and 853. A total of 5 individuals from families 19, 69, 273, and 853 did not carry the ARVD5-reference allele, 4, for marker *D3S3038*. Finally, the ARVD5-reference allele 4 for marker *D3S3659* was not observed in a total of 4 individuals from families 840 and 853.

It is worth noting that each marker was homozygous in some affected individuals. Given that ARVC in the NL population displays autosomal dominant inheritance, the marker homozygosity is likely due to the common occurrence of the designated disease-associated alleles in the general population.

By comparing genotypes of 21 individuals in 12 different families to the ARVD5-reference haplotype, the original 9.93 Mb candidate region was reduced to 2.36 Mb. The boundary markers for the ARVD5 were *D3S3610* and *D3S3613*. It is worth noting that the ARVD5-reference genotype for marker *D3S1516* was not shared in family 840 and is bordered by markers included in the narrowed candidate region. Excluding marker *D3S1516* would reduce the candidate region by only 475 bp. Therefore, to be conservative, we included the genomic region between markers *D3S2403* and *D3S3608* in further analysis. This significantly reduce the minimal critical region of the putative critical region by 7.57 Mb and excluded 30 of the 50 positional candidate genes. This newly defined 2.36 Mb region is predicted to contain the ARVD5 gene (Figure 3B).

3.1.1.1 Additional NL ARVC Family Members

During the course of this study, more patients were recruited and accessed by the clinical team. Subsequently, additional genotyping was carried out on individuals from thirteen of the fifteen previously investigated ARVD5-linked families. These included fifty-six additional ARVC clinically-affected patients and one hundred forty-five family members, with varying clinical signs. In addition, sixteen family members from five new families presenting with ARVC clinical signs were later genotyped. Haplotypes were constructed within existing family pedigrees (see Appendix D and Appendix E- data not shown for new families) to determine if additional informative recombinations were present. No additional recombination events were discovered within the ARVC clinically-affected

patients. Hence, the region could not be reduced beyond 2.36 Mb with the patient resources available to the project.

3.2 IDENTIFYING THE ARVD5 GENE

The majority of genes that cause ARVC code for for proteins that are involved in cell-cell adhesion and calcium channels (G. Thiene et al., 2007). We identified 20 annotated genes within the 2.36 Mb ARVD5 candidate region (Table 5 and Figure 3). The candidate genes were prioritized according to function and expression patterns. Candidate genes that are expressed in the heart, as well as those that had a desmosomal-related function were prioritized for screening.

Table 5: All genes annotated within the ARVD5 2.36 Mb region.
(As per) The Human Genome March 2006 Assembly, (Build 36.1). Exons include those for all isoforms. The amplicons cover all exons, untranslated regions, and their flanking sequences.

Gene Name	Gene Accession Number	Genomic Position		Strand	Genomic Size (bp)	No. of Exons	No. of Amplicons
		Start	End				
IQSEC1	NM_014869	13,003,536	12,917,079	–	86,457	13	17
NUP210	NM_024923	13,436,809	13,332,737	–	104,072	40	39
HDAC11	NM_024827	13,496,824	13,521,834	+	25,010	10	11
FBLN2	NM_001004019	13,565,625	13,654,922	+	89,297	18	27
WNT7A	NM_004625	13,896,619	13,835,083	–	61,536	4	11
TPRXL	AK092426	13,953,902	14,082,480	+	128,578	3	3
CHCHD4	NM_144636	14,141,323	14,128,584	–	12,739	4	4
TMEM43	NM_024334	14,141,546	14,160,180	+	18,634	12	16
XPC	NM_004628	14,195,143	14,161,651	–	33,492	16	18
LSM3	NM_014463	14,195,341	14,214,840	+	19,499	4	4
SLC6A6	NM_003043	14,419,110	14,503,973	+	84,863	15	28
GRIP2	NM_001080423	14,558,592	14,510,177	–	48,415	25	23
C3orf19	NM_016474	14,668,278	14,689,167	+	20,889	11	15
C3orf20	NM_032137	14,691,658	14,789,544	+	97,886	17	17
FGD5	NM_152536	14,835,810	14,950,899	+	115,089	20	23
NR2C2	NM_003298	14,964,240	15,065,782	+	101,542	15	29
MRPS25	NM_022497	15,081,820	15,065,024	–	16,796	4	11
ZFYVE20	NM_022340	15,115,659	15,086,584	–	29,075	14	23
CAPN7	NM_014296	15,222,737	15,269,426	+	46,689	21	25
SH3BP5	NM_004844*	15,349,108	15,271,250	–	77,858	10	12
					Total	**276**	**356**

A

Marker		D3S3610	D3S2403	D3S1516	D3S3608	D3S2385	D3S3602	D3S1585	D3S1554	D3S3595	D3S3613	D3S3473	D3S2338	D3S4547	D3S3510	D3S1293	D3S3038	D3S3669
Build 36.1 (Mb)		12.95	13.15	13.63	13.68	13.85	13.90	13.92	14.34	14.62	15.34	16.50	16.82	17.93	19.07	21.90	21.92	22.91
	932	246	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
	1139	nd	251	347	165	146	117	118	129	265	193	214	185	233	285	nd	nd	nd
	19	246	251	347	165	146	117	118	129	265	193	214	185	233	285	5 / 13	5	4
	64	246	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
	69	246	251	347	165	146	117	118	129	265	187	218	185	225 / 229	283	6 / 12	5 / 7	4
	76	256	251	347	165	146	117	118	129	265	193	214	185	233	285	7	4	4
	185	246	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
Family No.	273	246	251	347	165	146	117	118	129	265	195	214	185	229	283	8	5 / 6	4
	453	256	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
	581	246	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
	840	246	251	351	165	146	117	118	129	265	193	214	181	229	283	2 / 7	4	5
	853	246	251	347	165	146	117	118	129	265	193	214	185	233	285	12	5	3 / 5
	864	246	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
	964	246	251	347	165	146	117	118	129	265	193	214	185	233	285	8	4	4
	977	246	251	nd	165	146	117	118	129	265	193	nd	nd	233	nd	nd	nd	nd

B

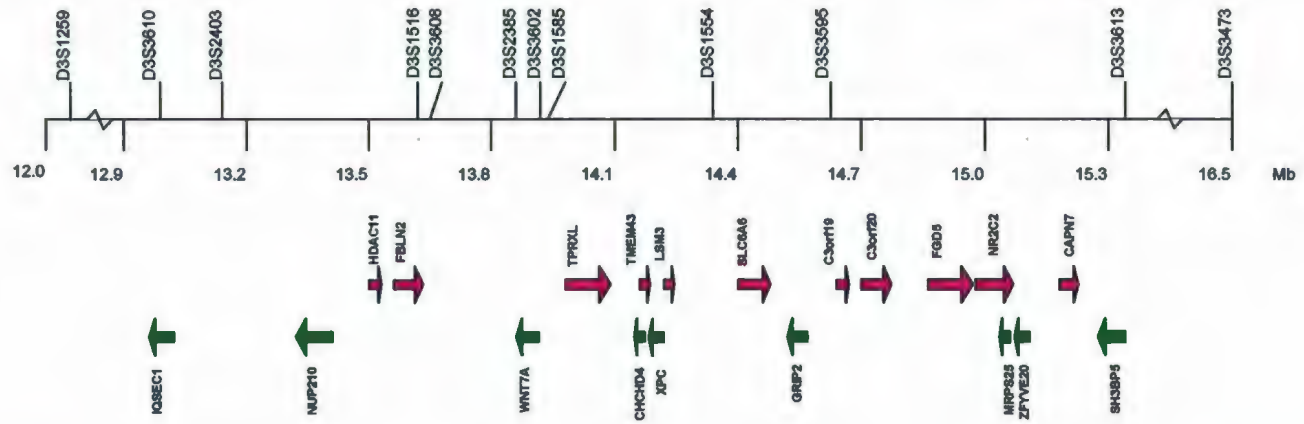


Figure 3:

A – Summary of the genotypes in ARVC clinically-affected patients from 15 ARVC families with NL ancestry. Yellow genotype blocks are consistent with the ARVD5-reference haplotype from families 932 and 1139. White genotype blocks are not consistent with the ARVD5-reference haplotype.
B – Physical map of the ARVD5 region on chromosome 3p25. Polymorphic markers and annotated genes are displayed on the map. Arrows indicate the direction of transcription of each positional candidate gene.

3.2.1 Sequence Variation Analysis

For the 20 ARVD5 positional candidate genes (i.e. *IQSEC1, NUP210, HDAC11, FBLN2, WNT7A, TPRXL, CHCHD4, TMEM43, XPC, LSM3, SLC6A6, GRIP2, C3orf19, C3orf20, FGD5, NR2C2, MRPS25, ZFYVE20, CAPN7* and *SH3BP5*) (Table 5), all of the untranslated regions (UTRs), exons and intron-exon boundaries were bidirectionally sequenced. A total of 356 amplicons were required to cover for the 276 exons in the 20 genes (Appendix 1).

The sequencing panel consisted of 4 affected family members and 3 unaffected controls - from families 64, 69, 453, and 840 - revealed a total of 240 sequence variations in 19 genes. No sequence alterations were found in the *LSM3* gene, thus excluding it from further analysis. Of the 240 variations, 19 variants in 8 genes (*HDAC11, TMEM43, XPC, SLC6A6, FGD5, NR2C2, MRPS25* and *CAPN7*) were found exclusively in the 4 ARVC affected family members (Table 6). This excluded genes *IQSEC1, NUP210, FBLN2, WNT7A, TPRXL, CHCHD4, GRIP2, C3orf19, C3orf20, ZFYVE20* and *SH3BP5* from further analysis.

Segregation analysis of the 19 variants found exclusively in the 4 affected patients was performed using DNA from family 1139. Eleven of these variants (highlighted in Table 6, Figure 4) segregated with the ARVD5-reference haplotype.

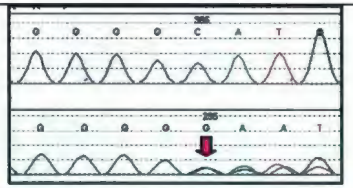
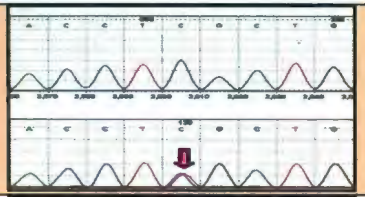
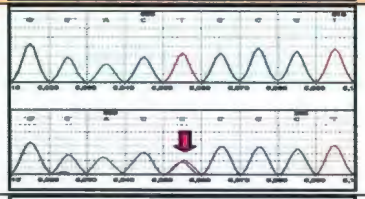
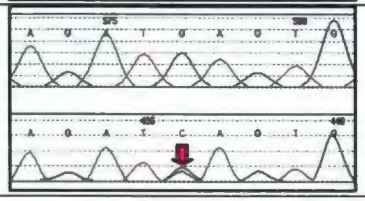
3.2.2 Using Population Controls to Determine Frequency

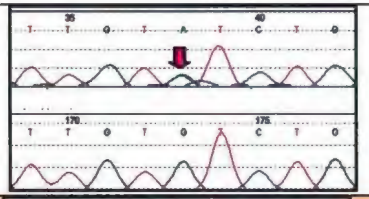
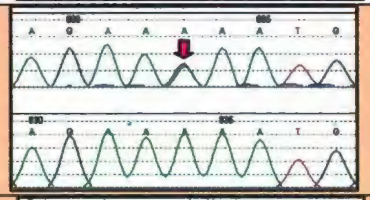
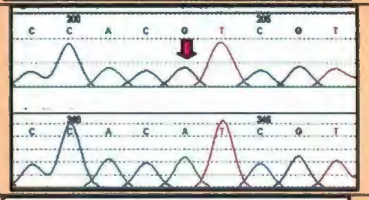
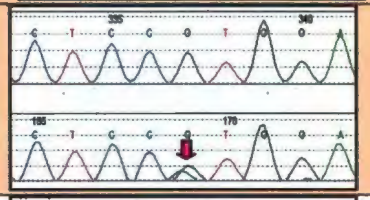
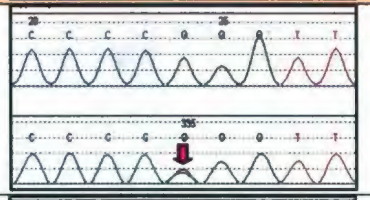
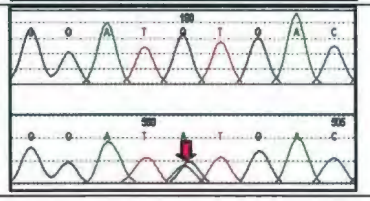
To determine if the 11 sequence variations segregating on the ARVD5-reference haplotype could be common DNA polymorphisms, we screened 44 to 161 control individuals from the NL population for each variation. Five of the 11 variants were rare sequence variants (< 1% of chromosomes tested) (Table 6).

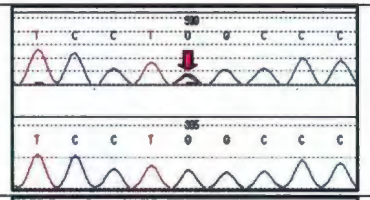
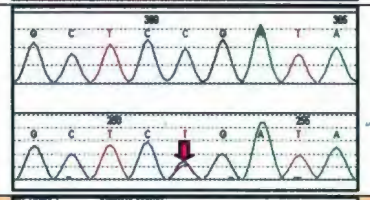
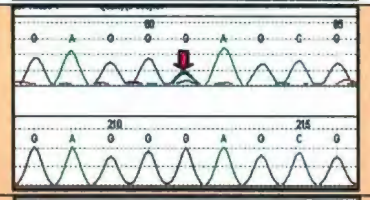
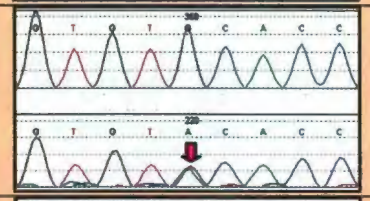
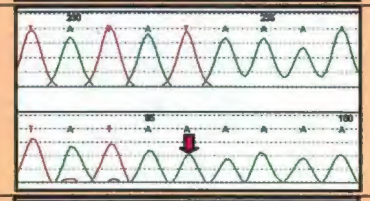
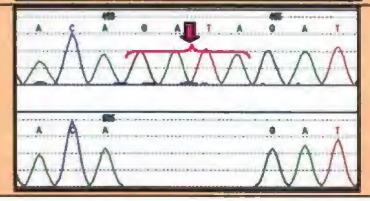
3.2.3 ARVC Gene Identification

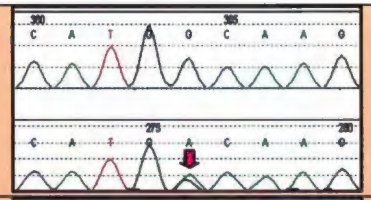
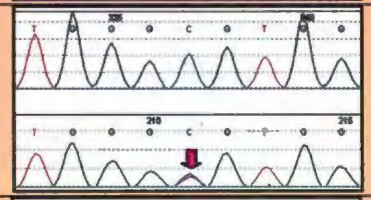
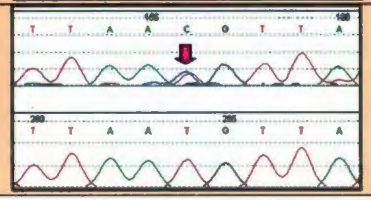
Segregation of the 5 rare variants was examined in families 69 and 273 (Figure 5). Only 1 variant, *TMEM43* (c.1073C>T), showed perfect co-segregation with the ARVC trait. All clinically affected and unaffected/unknown members of the ARVD5-linked families were then screened for the *TMEM43* (c.1073C>T) variant. All 83 ARVC clinically-affected patients, with available DNA samples, were positive for the genomic variation. One hundred fifty-one of the 212 patients who were clinical unaffected were negative for the genomic variation. The remaining 61 patients were mutation carriers. The unaffected family members who tested positive for the variant did not present with clinical signs at the time of clinical testing. Hence, these patients may be mutation carriers while being asymptomatic for the ARVC disease.

Table 6: Mutation screening results of the 20 positional candidate genes within the 2.39 Mb ARVD5 critical region. Nineteen variants are listed. They were found in 4 affected ARVC patients and absent in 3 unaffected relatives. The highlighted variants segregated with the ARVD5-reference haplotype.

Gene Name	Exon/Intron	**HGV Nomenclature**	A.A. Effect	Classification	Population Frequency per Chromosome	Electropherogram
HDAC11	Intron 4	c.369+18_369+19insG	None	Non-coding	ND	
TMEM43	Exon 12	c.1073C>T	S>L	Missense	0%	
TMEM43	3' UTR	c.1203+115T>C	None	Non-coding	ND	
XPC	3' UTR	c.2823+684G>C	None	Non-coding	ND	

SLC6A6	5' UTR	c.1-27420G>A	None	Non-coding	ND	
SLC6A6	Intron 5	c.599+370A>G	None	Non-coding	46%	
SLC6A6	Intron 6	c.733-1226A>G	None	Non-coding	55.6%	
FGD5	Exon 1	c.934G>A	V>M	Missense	0.6%	
FGD5	Intron 5	c.2186+22G>A	None	Non-coding	ND	
FGD5	Intron 5	c.2187-82G>A	None	Non-coding	ND	

FGD5	Exon 6	c.2220G>T	L>L	Synonymous	ND	
FGD5	Intron 10	c.2613+50C>T	None	Non-coding	ND	
FGD5	Intron 15	c.3085-74G>A	None	Non-coding	9.0%	
NR2C2	Intron 8	c.855+70G>A	None	Non-coding	9.1%	
NR2C2	3' UTR	c.1848+365T>A	None	Non-coding	17.8%	
NR2C2	3' UTR	c.1848+2965_1848 +2966insGATA	None	Non-coding	18.3%	

MRPS25	3' UTR	c.522+1059G>A	None	Non-coding	0%	
CAPN7	Intron 11	c.1289+68C>T	None	Non-coding	0.01%	
CAPN7	Intron 12	c.1430-28T>C	None	Non-coding	0%	

HGV = Human Genome Variation
A.A. = Amino Acid

Family 1139

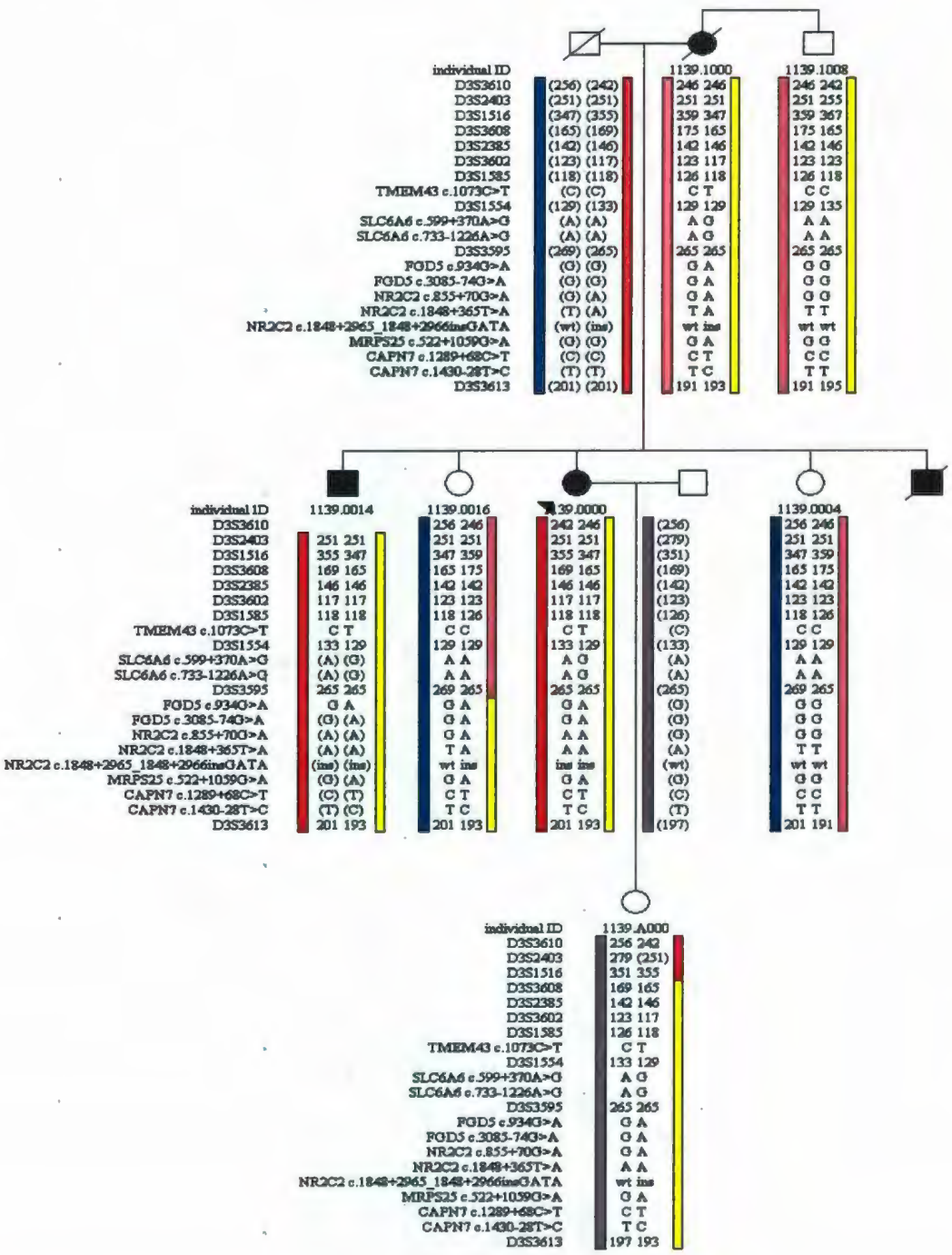


Figure 4: Pedigree with haplotypes of family 1139 demonstrating the segregation of 11 variants found exclusively in 4 affected NL ARVC patients from various other families.

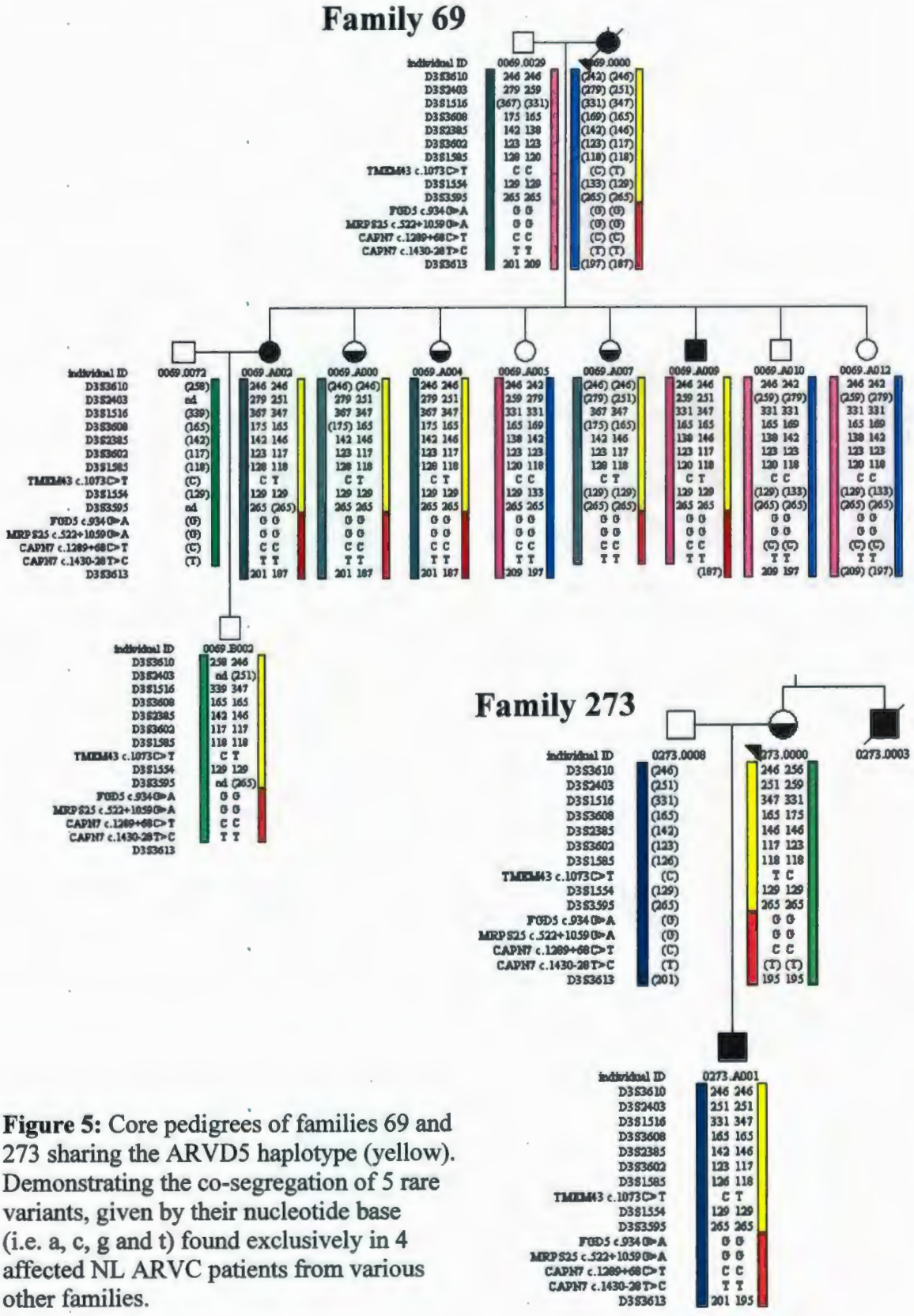


Figure 5: Core pedigrees of families 69 and 273 sharing the ARVD5 haplotype (yellow). Demonstrating the co-segregation of 5 rare variants, given by their nucleotide base (i.e. a, c, g and t) found exclusively in 4 affected NL ARVC patients from various other families.

CHAPTER 4: DISCUSSION

The primary aim of this research was to reduce the candidate disease region on 3p25 and to identify the gene causing ARVC at ARVD5. Fine mapping and haplotype analysis led to a important decrease in the ARVD5 region from a 9.93 Mb to 2.36 Mb. This reduction decreased the number of positional candidate genes from 58 to 20. This represents a 66% reduction in the number of genes that would require sequencing for mutation identification at ARVD5. Sequencing of the 20 annotated genes located in the 2.36 Mb region revealed one rare variant that was present in all ARVC clinically-affected patients, absent in population controls, and co-segregated with the ARVC trait.

Fine Mapping of 15 Families Reduces Critical Region

The strategy of defining a high resolution haplotype, common to all ARVC patients, is justified given our belief and probability that all families investigated originated from a common ancestor. High resolution haplotyping was achieved by adding 9 markers to those originally used to map the ARVD5 region (Ahmad et al., 1998) and initial genotyping of 33 ARVC clinically-diagnosed patients. These patients, from 15 families, shared a common haplotype between *D3S3610* and *D3S3613*, which was predicted to contain the ARVD5 gene. Because all clinically-affected patients and obligate carriers had the identical disease-associated haplotype, we suspect ARVC in these 15 NL families is caused by the same founder mutation.

Considering ARVC in the Newfoundland population at ARVD5 displays an autosomal dominant mode of inheritance, only one mutant allele is required for the expression of the disease. Occasionally, alleles were homozygous in ARVC clinically-affected individuals. This highlights the occurrence a marker's linked allele designation appearing in the general population. Assuming that a patient has one clinically-affected parent and one unaffected parent, a clinically-affected patient of an autosomal dominant disease who is homozygous with the linked allele suggests that one allele is linked to the diseased gene and the other is not. This showed a lack of informativeness of each marker due to the linked allele's designation appearing in the general population. This creates difficulties when tracing allelic transmission patterns. Also, as with all gene mapping studies, there still remains the possibility that this haplotype is shared by chance and do not harbor an disease-causing variant. However, with multiple clinically-affected families having several generations available for investigation and several large families available for genotyping and haplotyping, we are able to accurately represent the ARVD5-linked haplotype in patients.

Mutation Screening

Since ARVC in the Newfoundland population is considered a Mendelian disease we expected to observe a mutation leading to a functional change, such as a premature termination codon, a frameshift mutation or a missense mutation. The criteria used in this research to determine whether a genetic variant found can be considered a pathogenic variant takes into account several factors. The supporting evidence for the *TMEM43*

variant being disease-causing depends on the segregation analysis, the screening of 322 control chromosomes, and the conservation of genomic sequence surrounding the variant.

As a result of the sequencing analysis process, eleven variants were found to segregate on the ARVD5-associated haplotype in family R1139. Nine of these variants were in non-coding regions of the genome, presumably having little or no influence on the manufacture of proteins. Two variants were missense mutations. A missense mutation in exon 12 of *TMEM43* (c.1073C>T) changed the serine (polar, neutral and slightly hydrophilic) to a leucine (nonpolar, neutral and hydrophobic). The other missense mutation was located in exon 1 of *FGD5* (c.934G>A). This causes a substitution of the valine (nonpolar, neutral and hydrophobic) amino acid for methionine (nonpolar, neutral and hydrophobic). The segregation analysis was dependant on family size, the availability of patients' consent and the availability of DNA for our molecular genetic analysis.

Analysis of the population frequency determined that 5 of the 11 variants were rare sequence variants (< 1% of chromosomes tested) (Table 6). Further segregation analysis of the 5 rare variants in families 69 and 273 (Figure 5) showed that only the *TMEM43* (c.1073C>T) variant segregated perfectly with the ARVC clinically-affected members of these families and was absent in 322 control chromosomes. This large number of 322 control chromosomes were screened in order to increase the possibility of achieving a <1% polymorphism.

Bioinformatic evidence suggests that the wildtype serine amino acid at the *TMEM43* variant position is conserved throughout nine eukaryotic and prokaryotic species, with the exception of one type of bacteria (Merner et al., 2008).

The First Association of *TMEM43* with a Human Disease

The *TMEM43* discovery is an important one. This research illustrates the first association of this gene with a human trait or disease. Also, *TMEM43* is one of only three non-desmosomal genes proven to cause ARVC.

It is expected that many proteins have more than one function, depending on where they are found in the cell or within the body as a whole (DeSalle, American Museum of Natural History, & Yudell, 2005). Unfortunately, very little information had been previously published on the *TMEM43* gene. A small pool of data published to date can give speculation on the function of the *TMEM43* gene and how it relates to ARVC. Peroxisome proliferator-activated receptor gamma (PPARγ) is an adipogenic transcription factor that helps regulate fat formation (MIM 601487). A genome-wide scan for peroxisome proliferator response elements (PPREs) identified 1085 potential target genes of PPARγ, including *TMEM43*. This suggests that *TMEM43* may be a part of an adipogenic pathway regulated by PPARγ and may explain the fibrofatty replacement of the myocardium, a characteristic pathological finding in ARVC (Lemay & Hwang, 2006).

Our team carried out comparative sequence alignments of the *TMEM43* (c.1073C>T) variant region in mammalian, avian, amphibian, and insect orthologs. We found that the variant is fully conserved between these species.

Structurally, the *TMEM43* protein is predicted to be comprised of several different components. It contains four transmembrane domains. The variant occurs within the third predicted transmembrane domain of the *TMEM43* protein – imbedded in the phospholipid bilayer with no exposure to the cytoplasmic or extracellular environments. Four different bioinformatic programs predict the variant to be deleterious.

Eighteen phosphorylation sites and four YinOYang sites are located in the extracellular and cytoplasmic regions (Merner et al., 2008). A change in the phosphorylation state at one of these phosphorylation sites can change the function or localization of the protein (Peck, 2006). Another component of the *TMEM43* protein is an O-glycosylation site in the extracellular region. It is worth noting that the addition of N-acetyl-galactosamine to a serine or threonine residue at the O-glycosylation site, followed by carbohydrates, can alter the proteins' role in the formation of components for the extracellular matrix (Kozarsky, Kingsley, & Krieger, 1988; Peter-Katalinic, 2005). The transactivation domain and Small Ubiquitin-like Modifier (SUMO) attachment site are exclusively located in the cytoplasm region (Merner et al., 2008). The transactivation domain helps regulate gene expression. The SUMO attachment site can be used to alter the function of *TMEM43* by allowing SUMO proteins to attach and detach (Hay, 2005). Given previous

evidence of their function, several of these *TMEM43* proteins components can be suspect in causing the pathogenesis of ARVC at ARVD5.

TMEM43 is different from the desmosomal cadherins – desmocollin and desmoglein. Cadherins are structurally and functionally similar molecules (Ginsberg, DeSimone, & Geiger, 1991) that take part in selective calcium-dependent adhesion interactions between cell surfaces (Sonnenberg & Liem, 2007; Walsh et al., 1990). The desmosomal cadherins only have one transmembrane domain and the extracellular domain contains 4 repeats of ~110 residues each (Garrod, Merritt, & Nie, 2002; Ishii & Green, 2001). *TMEM43* contains four transmembrane domains and no extended residue repeats in either its extracellular or cytoplasmic regions. Protein sequence alignments of *TMEM43* with both desmocollin and desmoglein showed less than 10% similarity (Merner et al., 2008). Because *TMEM43* is structurally different from the desmosomal cadherins, deciphering the *TMEM43* biochemical pathways involved in the pathogenesis of ARVC will provide important insights to other ARVC genes.

CHAPTER 5: CONCLUSION & FUTURE PERSPECTIVES

We defined a critical region for the ARVD5 locus in the NL families. This region was found to contain 20 known genes. Mutation screening, using genomic DNA from 4 affected members and 3 unaffected controls, revealed 19 variants found exclusively in the affected members. Segregation analysis determined 11 of these polymorphisms segregated on the ARVD5-reference haplotype. Further segregation analysis of the 5 rare variants revealed that only *TMEM43* (c.1073C>T) had perfect segregation among the ARVC affected members of two ARVD5 linked families. This variation is a single heterozygous DNA alteration in exon 12 of *TMEM43*, which is predicted to change the serine encoded amino acid to a leucine. Based on our genetic evidence, we determined that *TMEM43* is the causative gene for ARVD5.

The interesting outcome of this research is the discovery of *TMEM43* as a new ARVC-causing gene. Although the function of *TMEM43* in fruit flies has been initiated by another member of our research team, the *in vivo* function of *TMEM43* is not fully understood. Functional studies may provide a new direction for identifying the biochemical pathways that are involved in the etiology of ARVC.

The *TMEM43* (c.1073C>T) mutation is considered the cause of ARVC in the Newfoundland population. Genetic testing is now available for this variant on a clinical basis and can be integrated into clinical practice as a first line test for ARVC. The implementation of a variant molecular diagnostic test has provided quick and reliable

diagnosis which will allow for closer monitoring and treatment being applied prophylactically. This will provide patients and their families with answers and, in some cases, peace of mind. There's also the potential for community health initiatives, such as population and age-specific screenings for this mutation.

To date, exonic sequence analysis of the known desmosomal ARVC-related genes has identified causative mutations in ~50% of ARVC probands (Awad et al., 2008). Given that the majority of the founding population of Newfoundland originated from Great Britain and Ireland it is expected that this mutation could be present in these and other populations around the world. Screening the remaining ~50% of patients for the *TMEM43* variant should be initiated.

Finding the underlying molecular pathology of *TMEM43* may provide the opportunity for identifying a drug treatment for the ARVC disorder. Further study of this variant is required to fully understand its molecular processes and its role in the occurrence of cardiomyopathies in humans.

REFERENCES

Ahmad, F., Li, D., Karibe, A., Gonzalez, O., Tapscott, T., Hill, R., et al. (1998). Localization of a gene responsible for arrhythmogenic right ventricular dysplasia to chromosome 3p23. *Circulation, 98*(25), 2791-2795.

Alcalai, R., Metzger, S., Rosenheck, S., Meiner, V., & Chajek-Shaul, T. (2003). A recessive mutation in desmoplakin causes arrhythmogenic right ventricular dysplasia, skin disorder, and woolly hair. *Journal of the American College of Cardiology, 42*(2), 319-327.

Andersson, L., Archibald, A., Ashburner, M., Audun, S., Barendse, W., Bitgood, J., et al. (1996). Comparative genome organization of vertebrates. the first international workshop on comparative genome organization. *Mammalian genome : official journal of the International Mammalian Genome Society, 7*(10), 717-734.

Andersson, L., Haley, C. S., Ellegren, H., Knott, S. A., Johansson, M., Andersson, K., et al. (1994). Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science (New York, N.Y.), 263*(5154), 1771-1774.

Antoniades, L., Tsatsopoulou, A., Anastasakis, A., Syrris, P., Asimaki, A., Panagiotakos, D., et al. (2006). Arrhythmogenic right ventricular cardiomyopathy caused by deletions in plakophilin-2 and plakoglobin (naxos disease) in families from greece and cyprus: Genotype-phenotype relations, diagnostic features and prognosis. *European Heart Journal, 27*(18), 2208-2216. doi:10.1093/eurheartj/ehl184

Arimura, T., Hayashi, T., Terada, H., Lee, S. Y., Zhou, Q., Takahashi, M., et al. (2004). A Cypher/ZASP mutation associated with dilated cardiomyopathy alters the binding affinity to protein kinase C. *The Journal of Biological Chemistry, 279*(8), 6746-6752. doi:10.1074/jbc.M311849200

Asimaki, A., Syrris, P., Wichter, T., Matthias, P., Saffitz, J. E., & McKenna, W. J. (2007). A novel dominant mutation in plakoglobin causes arrhythmogenic right ventricular cardiomyopathy. *American Journal of Human Genetics, 81*(5), 964-973. doi:10.1086/521633

Augusciak-Duma, A., & Sieron, A. L. (2008). Molecular characteristics of leiomyoma uteri based on selected compounds of the extracellular matrix. [Molekularna charakterystyka guzow leiomyoma uteri na przykladzie wybranych skladnikow macierzy pozakomorkowej] *Postepy Higieny i Medycyny Doswiadczalnej (Online), 62*, 148-165.

Avers, C. J. (1989). *Process and pattern in evolution.* New York: Oxford University Press.

Awad, M. M., Calkins, H., Judge, D. P., & Medscape. (2008). Mechanisms of disease: Molecular genetics of arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Nature Clinical Practice.Cardiovascular Medicine,* doi:10.1038/ncpcardio1182

Awad, M. M., Dalal, D., Tichnell, C., James, C., Tucker, A., Abraham, T., et al. (2006). Recessive arrhythmogenic right ventricular dysplasia due to novel cryptic splice mutation in PKP2. *Human Mutation, 27*(11), 1157. doi:10.1002/humu.9461

Basso, C., Czarnowska, E., Della Barbera, M., Bauce, B., Beffagna, G., Wlodarska, E. K., et al. (2006). Ultrastructural evidence of intercalated disc remodelling in

arrhythmogenic right ventricular cardiomyopathy: An electron microscopy investigation on endomyocardial biopsies. *European Heart Journal, 27*(15), 1847-1854. doi:10.1093/eurheartj/ehl095

Bauce, B., Basso, C., Rampazzo, A., Beffagna, G., Daliento, L., Frigo, G., et al. (2005). Clinical profile of four families with arrhythmogenic right ventricular cardiomyopathy caused by dominant desmoplakin mutations. *European Heart Journal, 26*(16), 1666-1675. doi:10.1093/eurheartj/ehi341

Bear, J. C., Nemec, T. F., Kennedy, J. C., Marshall, W. H., Power, A. A., Kolonel, V. M., et al. (1987). Persistent genetic isolation in outport newfoundland. *American Journal of Medical Genetics, 27*(4), 807-830. doi:10.1002/ajmg.1320270410

Bear, J. C., Nemec, T. F., Kennedy, J. C., Marshall, W. H., Power, A. A., Kolonel, V. M., et al. (1988). Inbreeding in outport newfoundland. *American Journal of Medical Genetics, 29*(3), 649-660. doi:10.1002/ajmg.1320290324

Beffagna, G., De Bortoli, M., Nava, A., Salamon, M., Lorenzon, A., Zaccolo, M., et al. (2007). Missense mutations in desmocollin-2 N-terminus, associated with arrhythmogenic right ventricular cardiomyopathy, affect intracellular localization of desmocollin-2 in vitro. *BMC Medical Genetics, 8*, 65. doi:10.1186/1471-2350-8-65

Beffagna, G., Occhi, G., Nava, A., Vitiello, L., Ditadi, A., Basso, C., et al. (2005). Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. *Cardiovascular Research, 65*(2), 366-373. doi:10.1016/j.cardiores.2004.10.005

Bengtsson, L., & Otto, H. (2008). LUMA interacts with emerin and influences its distribution at the inner nuclear membrane. *Journal of Cell Science, 121*(Pt 4), 536-548. doi:10.1242/jcs.019281

Bienengraeber, M., Olson, T. M., Selivanov, V. A., Kathmann, E. C., O'Cochlain, F., Gao, F., et al. (2004). ABCC9 mutations identified in human dilated cardiomyopathy disrupt catalytic KATP channel gating. *Nature Genetics, 36*(4), 382-387. doi:10.1038/ng1329

Bomma, C., Rutberg, J., Tandri, H., Nasir, K., Roguin, A., Tichnell, C., et al. (2004). Misdiagnosis of arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Journal of Cardiovascular Electrophysiology, 15*(3), 300-306. doi:10.1046/j.1540-8167.2004.03429.x

Bonne, G., Carrier, L., Richard, P., Hainque, B., & Schwartz, K. (1998). Familial hypertrophic cardiomyopathy: From mutations to functional defects. *Circulation Research, 83*(6), 580-593.

Botstein, D., & Risch, N. (2003). Discovering genotypes underlying human phenotypes: Past successes for mendelian disease, future approaches for complex disease. *Nature Genetics, 33 Suppl*, 228-237. doi:10.1038/ng1090

Botstein, D., White, R. L., Skolnick, M., & Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics, 32*(3), 314-331.

Bourgain, C., & Genin, E. (2005). Complex trait mapping in isolated populations: Are specific statistical methods required? *European Journal of Human Genetics : EJHG, 13*(6), 698-706. doi:10.1038/sj.ejhg.5201400

Chen, S. C., Balfour, I. C., & Jureidini, S. (2001). Clinical spectrum of restrictive cardiomyopathy in children. *The Journal of Heart and Lung Transplantation : The Official Publication of the International Society for Heart Transplantation, 20*(1), 90-92.

Choo, K. H. (1998). Why is the centromere so cold? *Genome Research, 8*(2), 81-82.

Cohen, S. N., Chang, A. C., Boyer, H. W., & Helling, R. B. (1973). Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences of the United States of America, 70*(11), 3240-3244.

Coonar, A. S., Protonotarios, N., Tsatsopoulou, A., Needham, E. W., Houlston, R. S., Cliff, S., et al. (1998). Gene for arrhythmogenic right ventricular cardiomyopathy with diffuse nonepidermolytic palmoplantar keratoderma and woolly hair (naxos disease) maps to 17q21. *Circulation, 97*(20), 2049-2058.

Corrado, D., Basso, C., Thiene, G., McKenna, W. J., Davies, M. J., Fontaliran, F., et al. (1997). Spectrum of clinicopathologic manifestations of arrhythmogenic right ventricular cardiomyopathy/dysplasia: A multicenter study. *Journal of the American College of Cardiology, 30*(6), 1512-1520.

Corrado, D., Fontaine, G., Marcus, F. I., McKenna, W. J., Nava, A., Thiene, G., et al. (2000). Arrhythmogenic right ventricular dysplasia/cardiomyopathy: Need for an international registry. european society of cardiology and the scientific council on cardiomyopathies of the world heart federation. *Journal of Cardiovascular Electrophysiology, 11*(7), 827-832.

Dalal, D., Molin, L. H., Piccini, J., Tichnell, C., James, C., Bomma, C., et al. (2006). Clinical features of arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in plakophilin-2. *Circulation, 113*(13), 1641-1649. doi:10.1161/CIRCULATIONAHA.105.568642

Denfield, S. W., Rosenthal, G., Gajarski, R. J., Bricker, J. T., Schowengerdt, K. O., Price, J. K., et al. (1997). Restrictive cardiomyopathies in childhood. etiologies and natural history. *Texas Heart Institute Journal / from the Texas Heart Institute of St.Luke's Episcopal Hospital, Texas Children's Hospital, 24*(1), 38-44.

DeSalle, R., American Museum of Natural History, & Yudell, M. (2005). *Welcome to the genome : A user's guide to the genetic past, present, and future*. Hoboken, NJ: Wiley-Liss.

Du, F. X., Woodward, B. W., & Denise, S. K. (1998). Haplotype construction of sires with progeny genotypes based on an exact likelihood. *Journal of Dairy Science, 81*(5), 1462-1468.

Evans, D. M., & Cardon, L. R. (2004). Guidelines for genotyping in genomewide linkage studies: Single-nucleotide-polymorphism maps versus microsatellite maps. *American Journal of Human Genetics, 75*(4), 687-692. doi:10.1086/424696

Garcia-Gras, E., Lombardi, R., Giocondo, M. J., Willerson, J. T., Schneider, M. D., Khoury, D. S., et al. (2006). Suppression of canonical Wnt/beta-catenin signaling by nuclear plakoglobin recapitulates phenotype of arrhythmogenic right ventricular cardiomyopathy. *The Journal of Clinical Investigation, 116*(7), 2012-2021. doi:10.1172/JCI27751

Garrod, D. R., Merritt, A. J., & Nie, Z. (2002). Desmosomal adhesion: Structural basis, molecular mechanism and regulation (review). *Molecular Membrane Biology, 19*(2), 81-94. doi:10.1080/09687680210132476

Gerull, B., Gramlich, M., Atherton, J., McNabb, M., Trombitas, K., Sasse-Klaassen, S., et al. (2002). Mutations of TTN, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nature Genetics, 30*(2), 201-204. doi:10.1038/ng815

Gerull, B., Heuser, A., Wichter, T., Paul, M., Basson, C. T., McDermott, D. A., et al. (2004). Mutations in the desmosomal protein plakophilin-2 are common in arrhythmogenic right ventricular cardiomyopathy. *Nature Genetics, 36*(11), 1162-1164. doi:10.1038/ng1461

Gerull, B., Osterziel, K. J., Witt, C., Dietz, R., & Thierfelder, L. (1998). A rapid protocol for cardiac troponin T gene mutation detection in familial hypertrophic cardiomyopathy. *Human mutation, 11*(2), 179-182. doi:2-W

Gewillig, M., Mertens, L., Moerman, P., & Dumoulin, M. (1996). Idiopathic restrictive cardiomyopathy in childhood. A diastolic disorder characterized by delayed relaxation. *European Heart Journal, 17*(9), 1413-1420.

Ginsberg, D., DeSimone, D., & Geiger, B. (1991). Expression of a novel cadherin (EP-cadherin) in unfertilized eggs and early xenopus embryos. *Development (Cambridge, England), 111*(2), 315-325.

Grant, G. R., Manduchi, E., Cheung, V. G., & Ewens, W. J. (1999). Significance testing for direct identity-by-descent mapping. *Annals of Human Genetics, 63*(Pt 5), 441-454.

Green, R. C., Green, J. S., Buehler, S. K., Robb, J. D., Daftary, D., Gallinger, S., et al. (2007). Very high incidence of familial colorectal cancer in newfoundland: A comparison with ontario and 13 other population-based studies. *Familial Cancer, 6*(1), 53-62. doi:10.1007/s10689-006-9104-x

Guo, S. W. (1994). Computation of identity-by-descent proportions shared by two siblings. *American Journal of Human Genetics, 54*(6), 1104-1109.

Haines, J. L., & Pericak-Vance, M. A. (2006). *Genetic analysis of complex diseases* (2nd ed.). Hoboken, N.J.: Wiley-Liss. from http://www.loc.gov/catdir/enhancements/fy0622/2005022234-d.html

Hamid, M. S., Norman, M., Quraishi, A., Firoozi, S., Thaman, R., Gimeno, J. R., et al. (2002). Prospective evaluation of relatives for familial arrhythmogenic right ventricular cardiomyopathy/dysplasia reveals a need to broaden diagnostic criteria. *Journal of the American College of Cardiology, 40*(8), 1445-1450.

Hay, R. T. (2005). SUMO: A history of modification. *Molecular Cell, 18*(1), 1-12. doi:10.1016/j.molcel.2005.03.012

Henke, W., Herdel, K., Jung, K., Schnorr, D., & Loening, S. A. (1997). Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Research, 25*(19), 3957-3958.

HERSHEY, A. D. (1952). Inheritance in bacteriophage. *Annals of the New York Academy of Sciences, 54*(6), 960-962.

Heuser, A., Plovie, E. R., Ellinor, P. T., Grossmann, K. S., Shin, J. T., Wichter, T., et al. (2006). Mutant desmocollin-2 causes arrhythmogenic right ventricular

cardiomyopathy. *American Journal of Human Genetics, 79*(6), 1081-1088. doi:10.1086/509044

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

Hodgkinson, K. A., Parfrey, P. S., Bassett, A. S., Kupprion, C., Drenckhahn, J., Norman, M. W., et al. (2005). The impact of implantable cardioverter-defibrillator therapy on survival in autosomal-dominant arrhythmogenic right ventricular cardiomyopathy (ARVD5). *Journal of the American College of Cardiology, 45*(3), 400-408. doi:10.1016/j.jacc.2004.08.068

Hofmann, S. L., Goldstein, J. L., Orth, K., Moomaw, C. R., Slaughter, C. A., & Brown, M. S. (1989). Molecular cloning of a histidine-rich Ca2+-binding protein of sarcoplasmic reticulum that contains highly conserved repeated elements. *The Journal of Biological Chemistry, 264*(30), 18083-18090.

Hostetler, J. A. (1985). History and relevance of the hutterite population for genetic studies. *American Journal of Medical Genetics, 22*(3), 453-462. doi:10.1002/ajmg.1320220303

Houwen, R. H., Baharloo, S., Blankenship, K., Raeymaekers, P., Juyn, J., Sandkuijl, L. A., et al. (1994). Genome screening by searching for shared segments: Mapping a gene for benign recurrent intrahepatic cholestasis. *Nature Genetics, 8*(4), 380-386. doi:10.1038/ng1294-380

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature, 431*(7011), 931-945. doi:10.1038/nature03001

Ishii, K., & Green, K. J. (2001). Cadherin function: Breaking the barrier. *Current Biology : CB, 11*(14), R569-72.

Kamisago, M., Sharma, S. D., DePalma, S. R., Solomon, S., Sharma, P., McDonough, B., et al. (2000). Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *The New England Journal of Medicine, 343*(23), 1688-1696.

Kaurah, P., MacMillan, A., Boyd, N., Senz, J., De Luca, A., Chun, N., et al. (2007). Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. *JAMA : The Journal of the American Medical Association, 297*(21), 2360-2372. doi:10.1001/jama.297.21.2360

Keller, L. F., Jeffery, K. J., Arcese, P., Beaumont, M. A., Hochachka, W. M., Smith, J. N., et al. (2001). Immigration and the ephemerality of a natural population bottleneck: Evidence from molecular markers. *Proceedings.Biological Sciences / the Royal Society, 268*(1474), 1387-1394. doi:10.1098/rspb.2001.1607

Kirsch, L. R., Weinstock, D. J., Magid, M. S., Levin, A. R., & Gold, J. P. (1993). Treatment of presumed arrhythmogenic right ventricular dysplasia in an adolescent. *Chest, 104*(1), 298-300.

Kozarsky, K., Kingsley, D., & Krieger, M. (1988). Use of a mutant cell line to study the kinetics and function of O-linked glycosylation of low density lipoprotein receptors. *Proceedings of the National Academy of Sciences of the United States of America, 85*(12), 4335-4339.

Kruglyak, L. (1997). What is significant in whole-genome linkage disequilibrium studies? *American Journal of Human Genetics, 61*(4), 810-812.

Kwon, J. M., & Goate, A. M. (2000). The candidate gene approach. *Alcohol Research & Health : The Journal of the National Institute on Alcohol Abuse and Alcoholism, 24*(3), 164-168.

Lahtinen, A. M., Lehtonen, A., Kaartinen, M., Toivonen, L., Swan, H., Widen, E., et al. (2008). Plakophilin-2 missense mutations in arrhythmogenic right ventricular cardiomyopathy. *International Journal of Cardiology, 126*(1), 92-100. doi:10.1016/j.ijcard.2007.03.137

Lander, E. S., & Botstein, D. (1987). Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children. *Science (New York, N.Y.), 236*(4808), 1567-1570.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860-921. doi:10.1038/35057062

Laurie, D. A., & Hulten, M. A. (1985). Further studies on bivalent chiasma frequency in human males with normal karyotypes. *Annals of Human Genetics, 49*(Pt 3), 189-201.

Lee, H. G., Kang, H., Kim, D. H., & Park, W. J. (2001). Interaction of HRC (histidine-rich ca(2+)-binding protein) and triadin in the lumen of sarcoplasmic reticulum. *The Journal of Biological Chemistry, 276*(43), 39533-39538. doi:10.1074/jbc.M010664200

Lemay, D. G., & Hwang, D. H. (2006). Genome-wide identification of peroxisome proliferator response elements using integrated computational genomics. *Journal of Lipid Research, 47*(7), 1583-1587. doi:10.1194/jlr.M500504-JLR200

Lewis, A. B. (1992). Clinical profile and outcome of restrictive cardiomyopathy in children. *American Heart Journal, 123*(6), 1589-1593.

Li, D., Ahmad, F., Gardner, M. J., Weilbaecher, D., Hill, R., Karibe, A., et al. (2000). The locus of a novel gene responsible for arrhythmogenic right-ventricular dysplasia characterized by early onset and high penetrance maps to chromosome 10p12-p14. *American Journal of Human Genetics, 66*(1), 148-156.

Li, D., Gonzalez, O., Bachinski, L. L., & Roberts, R. (2000). Human protein tyrosine phosphatase-like gene: Expression profile, genomic structure, and mutation analysis in families with ARVD. *Gene, 256*(1-2), 237-243.

Lindstrom, L., Nylander, E., Larsson, H., & Wranne, B. (2005). Left ventricular involvement in arrhythmogenic right ventricular cardiomyopathy - a scintigraphic and echocardiographic study. *Clinical Physiology and Functional Imaging, 25*(3), 171-177. doi:10.1111/j.1475-097X.2005.00607.x

Liu, N., Rizzi, N., Boveri, L., & Priori, S. G. (2008). Ryanodine receptor and calsequestrin in arrhythmogenesis: What we have learnt from genetic diseases and transgenic mice. *Journal of Molecular and Cellular Cardiology,* doi:10.1016/j.yjmcc.2008.10.012

Lotz, B. P., & van der Meyden, C. H. (1985). Myotonic dystrophy. part I. A genealogical study in the northern transvaal. *South African Medical Journal = Suid-Afrikaanse Tydskrif Vir Geneeskunde, 67*(20), 812-814.

LURIA, S. E., FRASER, D. K., ADAMS, J. N., & BURROUS, J. W. (1958). Lysogenization, transduction and genetic recombination in bacteria. *Cold Spring Harbor Symposia on Quantitative Biology, 23*, 71-82.

Luty, J. A., Guo, Z., Willard, H. F., Ledbetter, D. H., Ledbetter, S., & Litt, M. (1990). Five polymorphic microsatellite VNTRs on the human X chromosome. *American Journal of Human Genetics, 46*(4), 776-783.

Lynn, A., Ashley, T., & Hassold, T. (2004). Variation in human meiotic recombination. *Annual Review of Genomics and Human Genetics, 5*, 317-349. doi:10.1146/annurev.genom.4.070802.110217

Lynn, A., Kashuk, C., Petersen, M. B., Bailey, J. A., Cox, D. R., Antonarakis, S. E., et al. (2000). Patterns of meiotic recombination on the long arm of human chromosome 21. *Genome Research, 10*(9), 1319-1332.

Lynn, A., Koehler, K. E., Judis, L., Chan, E. R., Cherry, J. P., Schwartz, S., et al. (2002). Covariation of synaptonemal complex length and mammalian meiotic exchange rates. *Science (New York, N.Y.), 296*(5576), 2222-2225. doi:10.1126/science.1071220

MacLean, C. J., Martin, R. B., Sham, P. C., Wang, H., Straub, R. E., & Kendler, K. S. (2000). The trimmed-haplotype test for linkage disequilibrium. *American Journal of Human Genetics, 66*(3), 1062-1075. doi:10.1086/302796

MacRae, C. A., Birchmeier, W., & Thierfelder, L. (2006). Arrhythmogenic right ventricular cardiomyopathy: Moving toward mechanism. *The Journal of Clinical Investigation, 116*(7), 1825-1828. doi:10.1172/JCI29174

MacRae,C.A.(2), Birchmeier,W., Thierfelder,L. Arrhythmogenic right ventricular cardiomyopathy: Moving toward mechanism. JOURNAL OF CLINICAL INVESTIGATION. ONLINE Copyright 2006 by American Society for Clinical Investigation. Reproduced with permission of American Society for Clinical Investigation in the format Dissertation via Copyright Clearance Center.

Mannion, J. J., & 20 Memorial University of Newfoundland. Institute of Social and Economic Research. (1977). *The peopling of newfoundland : Essays in historical geography*. St. John's, Nfld.: Institute of Social and Economic Research, Memorial University of Newfoundland.

Marian, A. J., Yu, Q. T., Mares, A.,Jr, Hill, R., Roberts, R., & Perryman, M. B. (1992). Detection of a new mutation in the beta-myosin heavy chain gene in an individual with hypertrophic cardiomyopathy. *The Journal of clinical investigation, 90*(6), 2156-2165.

Marx, S. O., Reiken, S., Hisamatsu, Y., Jayaraman, T., Burkhoff, D., Rosemblit, N., et al. (2000). PKA phosphorylation dissociates FKBP12.6 from the calcium release channel (ryanodine receptor): Defective regulation in failing hearts. *Cell, 101*(4), 365-376.

Matise, T. C., Chen, F., Chen, W., De La Vega, F. M., Hansen, M., He, C., et al. (2007). A second-generation combined linkage physical map of the human genome. *Genome research, 17*(12), 1783-1786. doi:10.1101/gr.7156307

Mayr, E. (1963). The new versus the classical in science. *Science (New York, N.Y.), 141*(3583), 765. doi:10.1126/science.141.3583.765

McKenna, W. J., Thiene, G., Nava, A., Fontaliran, F., Blomstrom-Lundqvist, C., Fontaine, G., et al. (1994). Diagnosis of arrhythmogenic right ventricular dysplasia/cardiomyopathy. task force of the working group myocardial and pericardial disease of the european society of cardiology and of the scientific council on cardiomyopathies of the international society and federation of cardiology. *British Heart Journal, 71*(3), 215-218.

McKoy, G., Protonotarios, N., Crosby, A., Tsatsopoulou, A., Anastasakis, A., Coonar, A., et al. (2000). Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (naxos disease). *Lancet, 355*(9221), 2119-2124.

McPeek, M. S. (2000). From mouse to human: Fine mapping of quantitative trait loci in a model organism. *Proceedings of the National Academy of Sciences of the United States of America, 97*(23), 12389-12390. doi:10.1073/pnas.240463597

Melberg, A., Oldfors, A., Blomstrom-Lundqvist, C., Stalberg, E., Carlsson, B., Larrson, E., et al. (1999). Autosomal dominant myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy linked to chromosome 10q. *Annals of Neurology, 46*(5), 684-692.

Morera, B., & Barrantes, R. (2004). Is the central valley of costa rica a genetic isolate? *Revista De Biologia Tropical, 52*(3), 629-644.

Merner, N. D., Hodgkinson, K. A., Haywood, A. F., Connors, S., French, V. M., Drenckhahn, J. D., et al. (2008). Arrhythmogenic right ventricular cardiomyopathy type 5 is a fully penetrant, lethal arrhythmic disorder caused by a missense mutation in the TMEM43 gene. *American Journal of Human Genetics, 82*(4), 809-821. doi:10.1016/j.ajhg.2008.01.010

Merriman, T. R., Eaves, I. A., Twells, R. C., Merriman, M. E., Danoy, P. A., Muxworthy, C. E., et al. (1998). Transmission of haplotypes of microsatellite markers rather than single marker alleles in the mapping of a putative type 1 diabetes susceptibility gene (IDDM6). *Human Molecular Genetics, 7*(3), 517-524.

Mestroni, L., Rocco, C., Gregori, D., Sinagra, G., Di Lenarda, A., Miocic, S., et al. (1999). Familial dilated cardiomyopathy: Evidence for genetic and phenotypic heterogeneity. heart muscle disease study group. *Journal of the American College of Cardiology, 34*(1), 181-190.

Morgan, T. H., Sturtevant, A. H., Muller, H. J., & Bridges, C. B. (1922; 1922). *The mechanism of mendelian heredity* (Rev. ed.). New York: Holt.

Mogensen, J., Kubo, T., Duque, M., Uribe, W., Shaw, A., Murphy, R., et al. (2003). Idiopathic restrictive cardiomyopathy is part of the clinical expression of cardiac troponin I mutations. *The Journal of clinical investigation, 111*(2), 209-216.

Moric-Janiszewska, E., & Markiewicz-Loskot, G. (2007). Review on the genetics of arrhythmogenic right ventricular dysplasia. *Europace : European Pacing, Arrhythmias, and Cardiac Electrophysiology : Journal of the Working Groups on Cardiac Pacing, Arrhythmias, and Cardiac Cellular Electrophysiology of the European Society of Cardiology, 9*(5), 259-266. doi:10.1093/europace/eum034

Morris, A. P., Whittaker, J. C., & Balding, D. J. (2000). Bayesian fine-scale mapping of disease loci, by hidden markov models. *American Journal of Human Genetics, 67*(1), 155-169. doi:10.1086/302956

Morris, A. P., Whittaker, J. C., & Balding, D. J. (2002). Fine-scale mapping of disease loci via shattered coalescent modeling of genealogies. *American Journal of Human Genetics, 70*(3), 686-707. doi:10.1086/339271

Moxley, R. T.,3rd, Meola, G., Udd, B., & Ricker, K. (2002). Report of the 84th ENMC workshop: PROMM (proximal myotonic myopathy) and other myotonic dystrophy-like syndromes: 2nd workshop. 13-15th october, 2000, loosdrecht, the netherlands. *Neuromuscular Disorders : NMD, 12*(3), 306-317.

Murphy, R. T., Mogensen, J., Shaw, A., Kubo, T., Hughes, S., & McKenna, W. J. (2004). Novel mutation in cardiac troponin I in recessive idiopathic dilated cardiomyopathy. *Lancet, 363*(9406), 371-372. doi:10.1016/S0140-6736(04)15468-8

Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science (New York, N.Y.), 310*(5746), 321-324. doi:10.1126/science.1117196

Nattel, S., & Schott, J. J. (2005). Arrhythmogenic right ventricular dysplasia type 1 and mutations in transforming growth factor beta3 gene regulatory regions: A breakthrough? *Cardiovascular Research, 65*(2), 302-304. doi:10.1016/j.cardiores.2004.11.023

Nava, A., Thiene, G., Canciani, B., Scognamiglio, R., Daliento, L., Buja, G., et al. (1988). Familial occurrence of right ventricular dysplasia: A study involving nine families. *Journal of the American College of Cardiology, 12*(5), 1222-1228.

Nolte, I. M., & Te Meerman, G. J. (2002). The probability that similar haplotypes are identical by descent. *Annals of Human Genetics, 66*(Pt 3), 195-209. doi:doi:10.1017/S0003480002001100

Norgett, E. E., Hatsell, S. J., Carvajal-Huerta, L., Cabezas, J. C., Common, J., Purkis, P. E., et al. (2000). Recessive mutation in desmoplakin disrupts desmoplakin-intermediate filament interactions and causes dilated cardiomyopathy, woolly hair and keratoderma. *Human Molecular Genetics, 9*(18), 2761-2766.

Norman, M., Simpson, M., Mogensen, J., Shaw, A., Hughes, S., Syrris, P., et al. (2005). Novel mutation in desmoplakin causes arrhythmogenic left ventricular cardiomyopathy. *Circulation, 112*(5), 636-642. doi:10.1161/CIRCULATIONAHA.104.532234

O'Brien, S. J., Menotti-Raymond, M., Murphy, W. J., Nash, W. G., Wienberg, J., Stanyon, R., et al. (1999). The promise of comparative genomics in mammals. *Science (New York, N.Y.), 286*(5439), 458-62, 479-81.

Olson, T. M., Illenberger, S., Kishimoto, N. Y., Huttelmaier, S., Keating, M. T., & Jockusch, B. M. (2002). Metavinculin mutations alter actin interaction in dilated cardiomyopathy. *Circulation, 105*(4), 431-437.

Olson, T. M., Kishimoto, N. Y., Whitby, F. G., & Michels, V. V. (2001). Mutations that alter the surface charge of alpha-tropomyosin are associated with dilated cardiomyopathy. *Journal of Molecular and Cellular Cardiology, 33*(4), 723-732. doi:10.1006/jmcc.2000.1339

Oxford, E. M., Musa, H., Maass, K., Coombs, W., Taffet, S. M., & Delmar, M. (2007). Connexin43 remodeling caused by inhibition of plakophilin-2 expression in cardiac cells. *Circulation Research, 101*(7), 703-711. doi:10.1161/CIRCRESAHA.107.154252

Peck, S. C. (2006). Analysis of protein phosphorylation: Methods and strategies for studying kinases and substrates. *The Plant Journal : For Cell and Molecular Biology, 45*(4), 512-522. doi:10.1111/j.1365-313X.2005.02613.x

Peltonen, L., Palotie, A., & Lange, K. (2000). Use of population isolates for mapping complex traits. *Nature Reviews.Genetics, 1*(3), 182-190. doi:10.1038/35042049

Peter-Katalinic, J. (2005). Methods in enzymology: O-glycosylation of proteins. *Methods in Enzymology, 405*, 139-171. doi:10.1016/S0076-6879(05)05007-X

Pilichou, K., Nava, A., Basso, C., Beffagna, G., Bauce, B., Lorenzon, A., et al. (2006). Mutations in desmoglein-2 gene are associated with arrhythmogenic right ventricular cardiomyopathy. *Circulation, 113*(9), 1171-1179. doi:10.1161/CIRCULATIONAHA.105.583674

Primrose, S. B., & Twyman, R. M. (2003). *Principles of genome analysis and genomics* (3rd ed.). Malden, Mass.: Blackwell Pub.

Quarrell, O. W., Rigby, A. S., Barron, L., Crow, Y., Dalton, A., Dennis, N., et al. (2007). Reduced penetrance alleles for huntington's disease: A multi-centre direct observational study. *Journal of Medical Genetics, 44*(3), e68. doi:10.1136/jmg.2006.045120

Rahman, P., Jones, A., Curtis, J., Bartlett, S., Peddle, L., Fernandez, B. A., et al. (2003). The newfoundland population: A unique resource for genetic investigation of complex diseases. *Human Molecular Genetics, 12 Spec No 2*, R167-72. doi:10.1093/hmg/ddg257

Rampazzo, A., Nava, A., Malacrida, S., Beffagna, G., Bauce, B., Rossi, V., et al. (2002). Mutation in human desmoplakin domain binding to plakoglobin causes a dominant form of arrhythmogenic right ventricular cardiomyopathy. *American Journal of Human Genetics, 71*(5), 1200-1206.

Rampazzo, A., Nava, A., Miorin, M., Fonderico, P., Pope, B., Tiso, N., et al. (1997). ARVD4, a new locus for arrhythmogenic right ventricular cardiomyopathy, maps to chromosome 2 long arm. *Genomics, 45*(2), 259-263. doi:10.1006/geno.1997.4927

Richardson, P., McKenna, W., Bristow, M., Maisch, B., Mautner, B., O'Connell, J., et al. (1996). Report of the 1995 world health Organization/International society and federation of cardiology task force on the definition and classification of cardiomyopathies. *Circulation, 93*(5), 841-842.

Ridley, M. (2004). *Evolution* (3rd ed.). Oxford: Blackwell Pub.

Riordan, J. R., Rommens, J. M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., et al. (1989). Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science (New York, N.Y.), 245*(4922), 1066-1073.

Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology (Clifton, N.J.), 132*, 365-386.

Salvatori, S., Fanin, M., Trevisan, C. P., Furlan, S., Reddy, S., Nagy, J. I., et al. (2005). Decreased expression of DMPK: Correlation with CTG repeat expansion and fibre

type composition in myotonic dystrophy type 1. *Neurological Sciences : Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 26*(4), 235-242. doi:10.1007/s10072-005-0466-x

Schmitt, J. P., Kamisago, M., Asahi, M., Li, G. H., Ahmad, F., Mende, U., et al. (2003). Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science (New York, N.Y.), 299*(5611), 1410-1413. doi:10.1126/science.1081578

Schonberger, J., Levy, H., Grunig, E., Sangwatanaroj, S., Fatkin, D., MacRae, C., et al. (2000). Dilated cardiomyopathy and sensorineural hearing loss: A heritable syndrome that maps to 6q23-24. *Circulation, 101*(15), 1812-1818.

Sen-Chowdhry, S., Lowe, M. D., Sporton, S. C., & McKenna, W. J. (2004). Arrhythmogenic right ventricular cardiomyopathy: Clinical presentation, diagnosis, and management. *The American Journal of Medicine, 117*(9), 685-695. doi:10.1016/j.amjmed.2004.04.028

Sen-Chowdhry, S., Syrris, P., & McKenna, W. J. (2005). Genetics of right ventricular cardiomyopathy. *Journal of Cardiovascular Electrophysiology, 16*(8), 927-935. doi:10.1111/j.1540-8167.2005.40842.x

Sen-Chowdhry, S., Syrris, P., Ward, D., Asimaki, A., Sevdalis, E., & McKenna, W. J. (2007). Clinical and genetic characterization of families with arrhythmogenic right ventricular dysplasia/cardiomyopathy provides novel insights into patterns of disease expression. *Circulation, 115*(13), 1710-1720. doi:10.1161/CIRCULATIONAHA.106.660241

Sen-Chowdhry, S., Syrris, P., & McKenna, W. J. (2007). Role of genetic analysis in the management of patients with arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Journal of the American College of Cardiology, 50*(19), 1813-1821. doi:10.1016/j.jacc.2007.08.008

Severini, G. M., Krajinovic, M., Pinamonti, B., Sinagra, G., Fioretti, P., Brunazzi, M. C., et al. (1996). A new locus for arrhythmogenic right ventricular dysplasia on the long arm of chromosome 14. *Genomics, 31*(2), 193-200.

Shi, Q., Spriggs, E., Field, L. L., Rademaker, A., Ko, E., Barclay, L., et al. (2002). Absence of age effect on meiotic recombination between human X and Y chromosomes. *American Journal of Human Genetics, 71*(2), 254-261. doi:10.1086/341559

Shifman, S., & Darvasi, A. (2001). The value of isolated populations. *Nature Genetics, 28*(4), 309-310. doi:10.1038/91060

Slatkin, M. (2008). Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nature Reviews.Genetics, 9*(6), 477-485. doi:10.1038/nrg2361

Song, K., Dube, M. P., Lim, J., Hwang, I., Lee, I., & Kim, J. J. (2007). Lamin A/C mutations associated with familial and sporadic cases of dilated cardiomyopathy in koreans. *Experimental & molecular medicine, 39*(1), 114-120.

Sonnenberg, A., & Liem, R. K. (2007). Plakins in development and disease. *Experimental Cell Research, 313*(10), 2189-2203. doi:10.1016/j.yexcr.2007.03.039

Syrris, P., Ward, D., Asimaki, A., Sen-Chowdhry, S., Ebrahim, H. Y., Evans, A., et al. (2006). Clinical expression of plakophilin-2 mutations in familial arrhythmogenic

right ventricular cardiomyopathy. *Circulation, 113*(3), 356-364. doi:10.1161/CIRCULATIONAHA.105.561654

Syrris, P., Ward, D., Evans, A., Asimaki, A., Gandjbakhch, E., Sen-Chowdhry, S., et al. (2006). Arrhythmogenic right ventricular dysplasia/cardiomyopathy associated with mutations in the desmosomal gene desmocollin-2. *American Journal of Human Genetics, 79*(5), 978-984. doi:10.1086/509122

Tateishi, H., Yano, M., Mochizuki, M., Suetomi, T., Ono, M., Xu, X., et al. (2008). Defective domain-domain interactions within the ryanodine receptor as a critical cause of diastolic Ca2+ leak in failing hearts. *Cardiovascular Research,* doi:10.1093/cvr/cvn303

Taylor, G. R., Day, I. N. M., & Human Genome Organisation. (2005). *Guide to mutation detection*. Hoboken, N.J.: Wiley. Retrieved from http://www.loc.gov/catdir/description/wiley042/2004004692.html; http://www.loc.gov/catdir/enhancements/fy0616/2004004692-b.html

te Meerman, G. J., & Van der Meulen, M. A. (1997). Genomic sharing surrounding alleles identical by descent: Effects of genetic drift and population growth. *Genetic Epidemiology, 14*(6), 1125-1130. doi:2-I

Thiene, G., Corrado, D., & Basso, C. (2007). Arrhythmogenic right ventricular cardiomyopathy/dysplasia. *Orphanet Journal of Rare Diseases, 2*, 45. doi:10.1186/1750-1172-2-45

Thiene, G., Nava, A., Corrado, D., Rossi, L., & Pennelli, N. (1988). Right ventricular cardiomyopathy and sudden death in young people. *The New England Journal of Medicine, 318*(3), 129-133.

Thiene, G., Basso, C., Danieli, G., Rampazzo, A., Corrado, D., & Nava, A. (1997). Arrhythmogenic right ventricular cardiomyopathy A still underrecognized clinic entity. *Trends in Cardiovascular Medicine, 7*(3), 84-90.

Tiso, N., Stephan, D. A., Nava, A., Bagattin, A., Devaney, J. M., Stanchi, F., et al. (2001). Identification of mutations in the cardiac ryanodine receptor gene in families affected with arrhythmogenic right ventricular cardiomyopathy type 2 (ARVD2). *Human Molecular Genetics, 10*(3), 189-194.

Turnpenny, P. D., Ellard, S., & Emery, A. E. H. (2007). *Emery's elements of medical genetics* (13th ed.). London: Elsevier/Churchill Livingstone.

Uzumcu, A., Norgett, E. E., Dindar, A., Uyguner, O., Nisli, K., Kayserili, H., et al. (2006). Loss of desmoplakin isoform I causes early onset cardiomyopathy and heart failure in a naxos-like syndrome. *Journal of Medical Genetics, 43*(2), e5. doi:10.1136/jmg.2005.032904

Van der Meulen, M. A., & te Meerman, G. J. (1997). Haplotype sharing analysis in affected individuals from nuclear families with at least one affected offspring. *Genetic Epidemiology, 14*(6), 915-920. doi:2-P

Van Tintelen, J. P., Entius, M. M., Bhuiyan, Z. A., Jongbloed, R., Wiesfeld, A. C., Wilde, A. A., et al. (2006). Plakophilin-2 mutations are the major determinant of familial arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circulation, 113*(13), 1650-1658. doi:10.1161/CIRCULATIONAHA.105.609719

Walsh, F. S., Barton, C. H., Putt, W., Moore, S. E., Kelsell, D., Spurr, N., et al. (1990). N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin gene. *Journal of Neurochemistry, 55*(3), 805-812.

Wang, T., Xue, N., & Birdwell, J. D. (2006). Least-square deconvolution: A framework for interpreting short tandem repeat mixtures. *Journal of Forensic Sciences, 51*(6), 1284-1297. doi:10.1111/j.1556-4029.2006.00268.x

Watson, J. D., & Watson, J. D. (2007). *Recombinant DNA : Genes and genomes--a short course* (3rd ed.). New York: W.H. Freeman.

Weissensteiner, T., & Lanchbury, J. S. (1996). Strategy for controlling preferential amplification and avoiding false negatives in PCR typing. *BioTechniques, 21*(6), 1102-1108.

Wichter, T., Paul, T. M., Eckardt, L., Gerdes, P., Kirchhof, P., Bocker, D., et al. (2005). Arrhythmogenic right ventricular cardiomyopathy. antiarrhythmic drugs, catheter ablation, or ICD? *Herz, 30*(2), 91-101. doi:10.1007/s00059-005-2677-6

Yang, Z., Bowles, N. E., Scherer, S. E., Taylor, M. D., Kearney, D. L., Ge, S., et al. (2006). Desmosomal dysfunction due to mutations in desmoplakin causes arrhythmogenic right ventricular dysplasia/cardiomyopathy. *Circulation Research, 99*(6), 646-655. doi:10.1161/01.RES.0000241482.19382.c6

Yu, J., Lazzeroni, L., Qin, J., Huang, M. M., Navidi, W., Erlich, H., et al. (1996). Individual variation in recombination among human males. *American Journal of Human Genetics, 59*(6), 1186-1192.

Zavattari, P., Deidda, E., Whalen, M., Lampis, R., Mulargia, A., Loddo, M., et al. (2000). Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: Demography, chromosome recombination frequency and selection. *Human Molecular Genetics, 9*(20), 2947-2957.

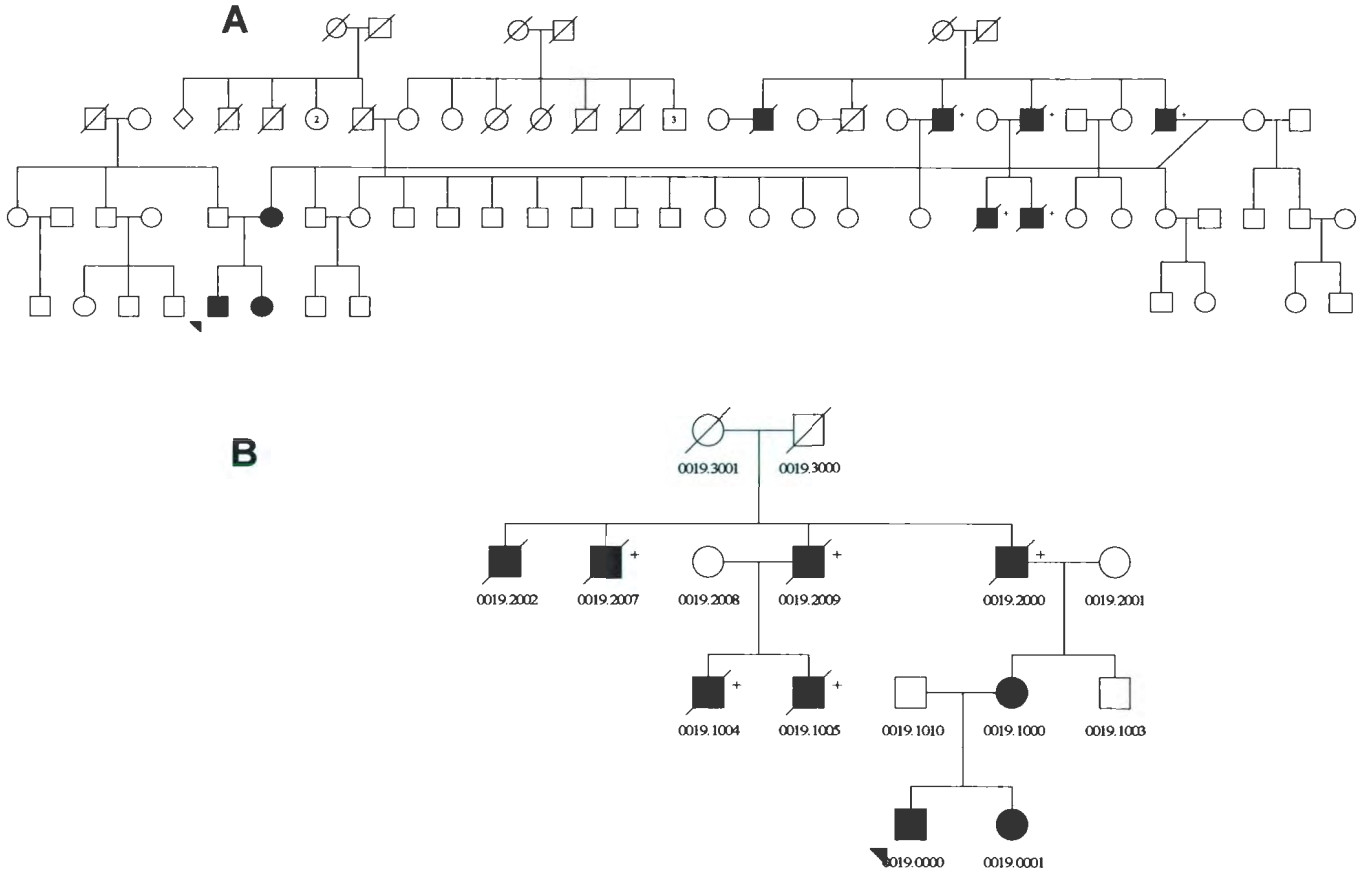
APPENDIX

Appendix A: PCR Product Purification, Sequencing and Ethanol Precipitation Protocols

PCR products were purified using Sephacryl S-300HR (supplied by GE Healthcare, Amersham Biosciences). The Sephacryl preparation procedure was as follows: poured off the supernatant ethanol solution from the Sephacryl; added 500-1000 ml TE buffer to the Sephacryl and gently inverted to re-suspend; Sephacryl was left to settle; gently poured off the supernatant; repeated TE buffer washes three more times; resuspended Sephacryl in an equal amount of TE buffer; stored at 4°C. The PCR purification procedure was as follows: re-suspended the prepared Sephacryl; added 300 µL of Sephacryl to each well in a Loprodyne membrane, 0.45µm pore, Nunc Silent Screen plate; plate spun 5 minutes at 3000 rpm; discarded waste; added PCR products to Silent Screen plate wells treated with Sephacryl; centrifuged 5 minutes at 3000 rpm; the flow-through contained purified PCR products.

Direct sequencing for both the forward and reverse stands was performed with a BigDye Terminator DNA sequencing kit (version 3.1) in an ABI MicroAMP PCR plate. Each reaction contained 0.5 µL of Sequencing Mix, 2 µL 5X Sequencing Buffer, 1.6 mM Primer, 0.5-2 µL purified DNA template, and autoclaved distilled water to a total volume of 20 µL. Amplification was performed on a 9700 GeneAmp PCR system (ABI Applied Biosystems) with the following conditions: denatured at 96°C for 1 min; 25 cycles of denaturing at 96°C for 30 sec, annealing at 50°C for 5 sec, and extension at 60°C for 4 min; cooled and maintained at 4°C.

The ethanol precipitation of sequencing reactions protocol was as follows: following PCR, 5 µL of 125 mM ethylenediamine tetraacetic acid (EDTA) and 65 µL of 95% ethanol was added to each reaction well; precipitated for 1-20 hours with no light source; centrifuged 30 min at 3000 x g; gently inverted to decant ethanol; added 150 µL of 70% ethanol; centrifuged 15 min at 3000 x g; gently inverted to decant ethanol; allowed the ethanol to evaporate completely for 10-20 minutes with no light source; added 30 µL deionised dimethylformamide to each well; gently vortexed to re-suspend; spun briefly; denatured 2 min at 95°C; and chilled to ~4°C.

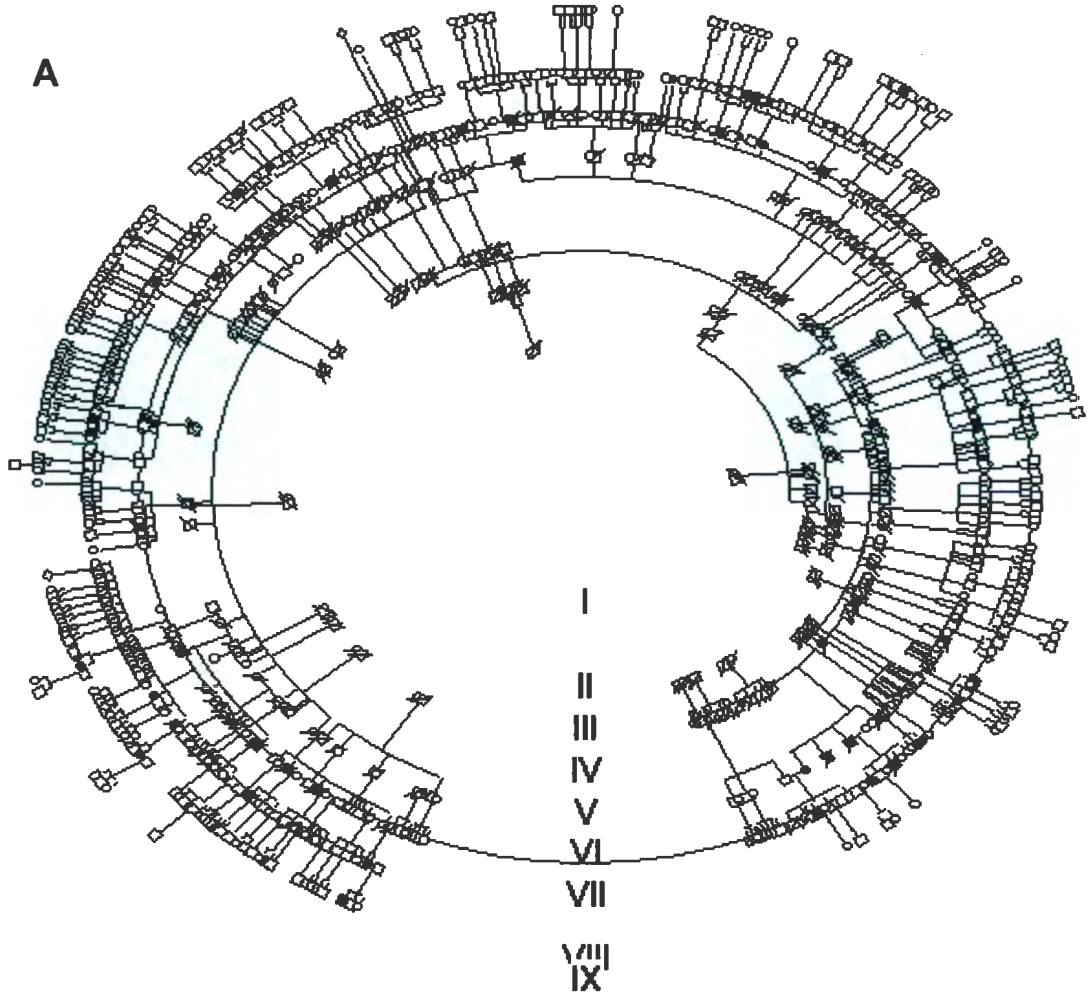


B1 – Family 19

Appendix B: Pedigree structures of 15 ARVC clinically affected NL families used for recombination mapping.

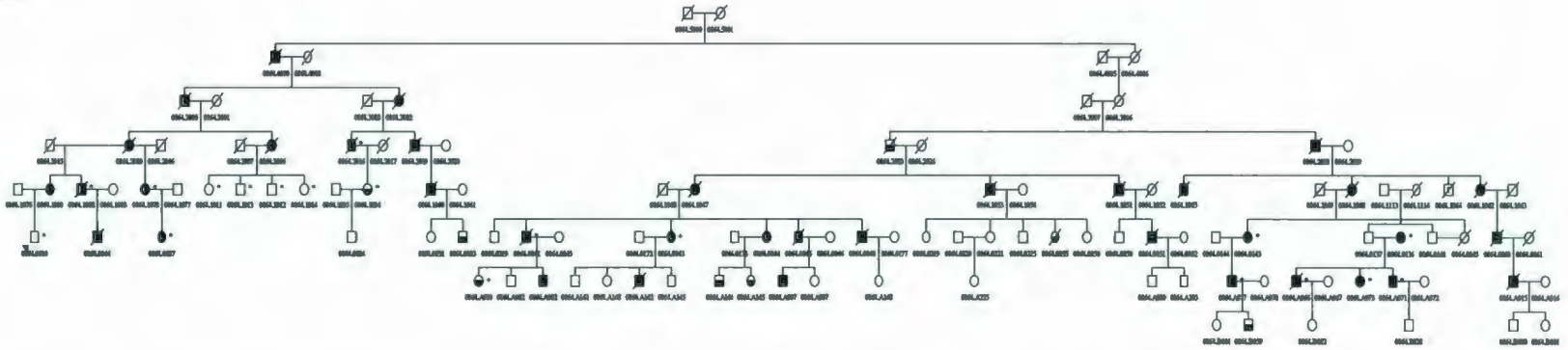
A - Complete pedigree structure.

B - Abbreviated pedigree structure used in the study.

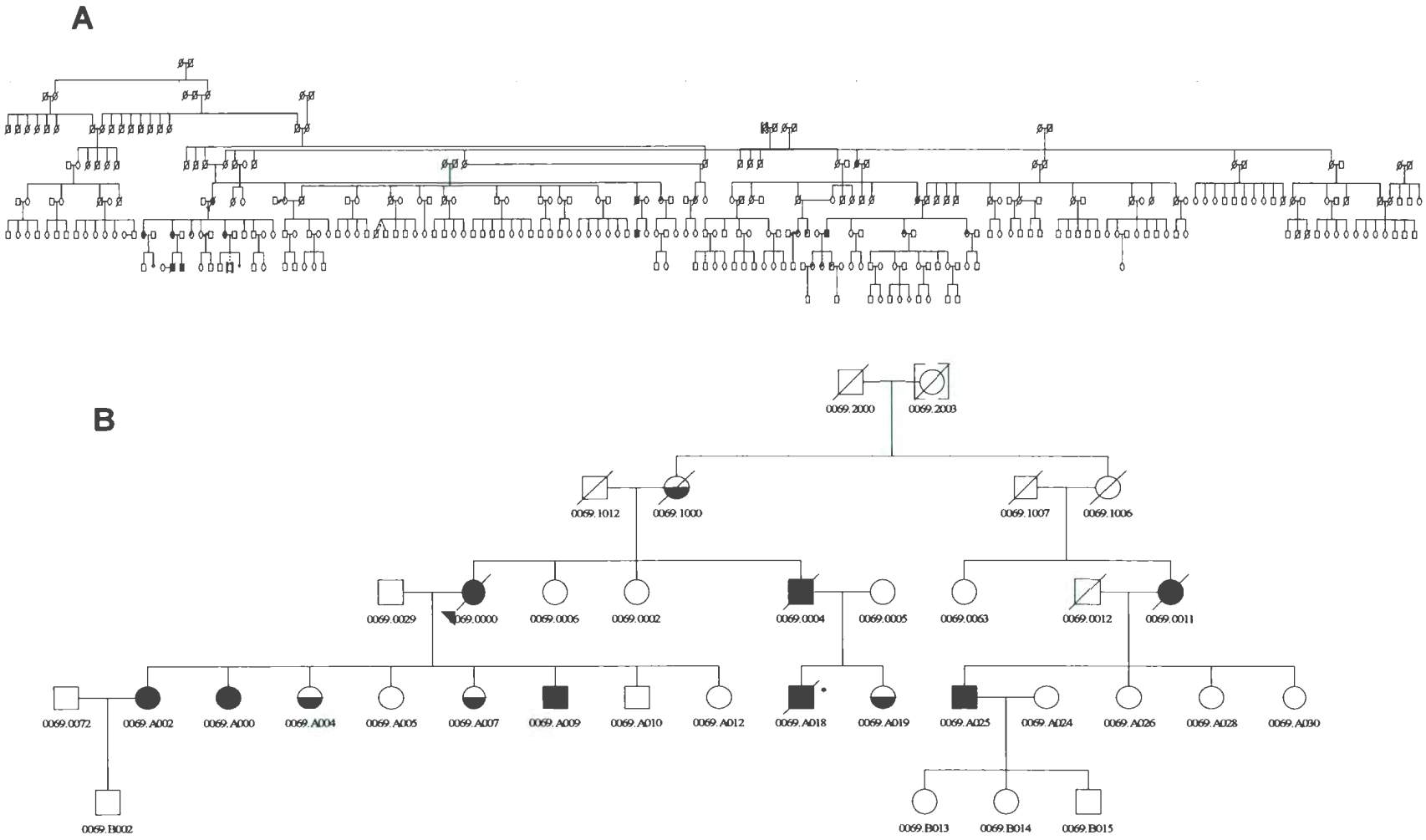


B2 – Family 64

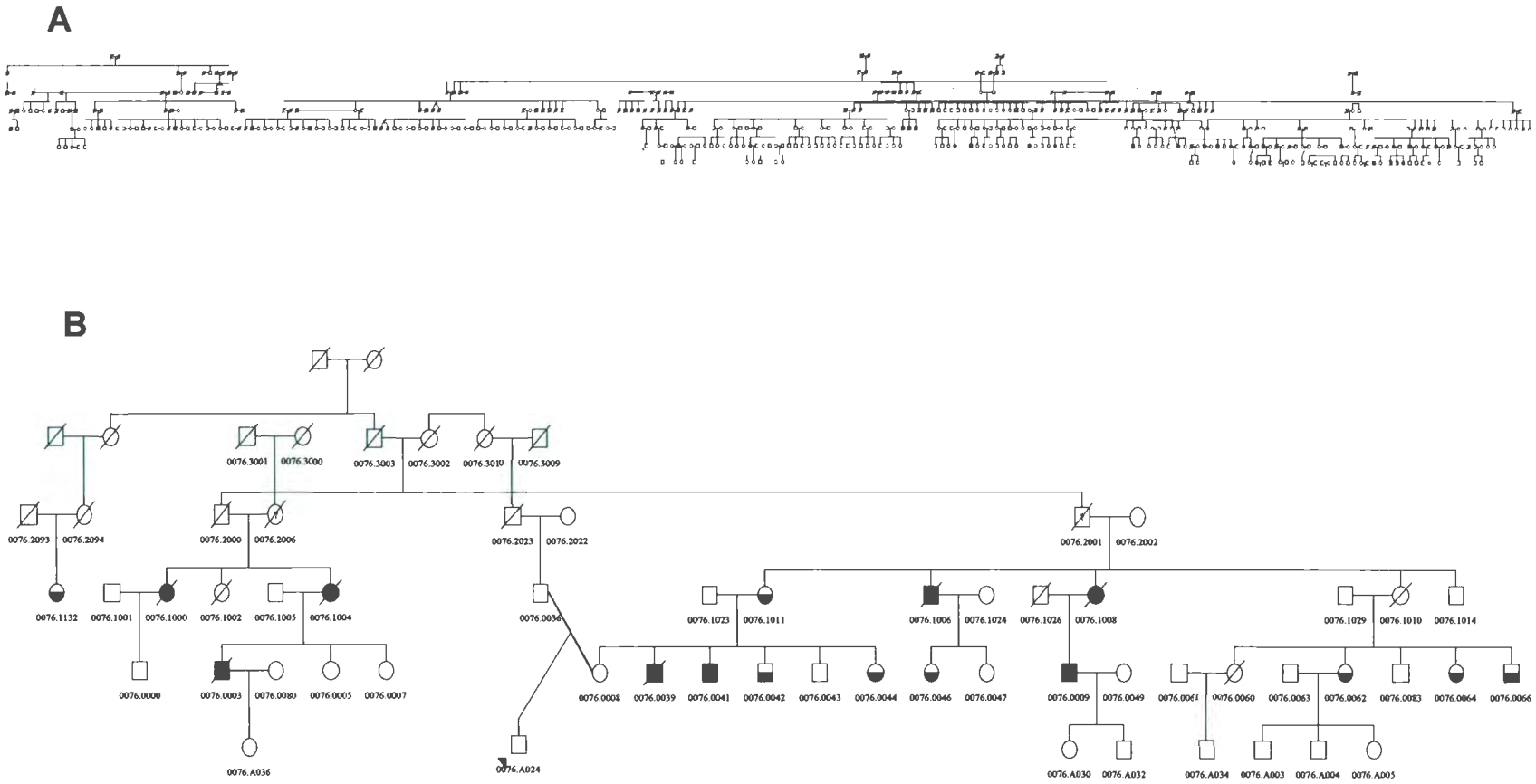
B



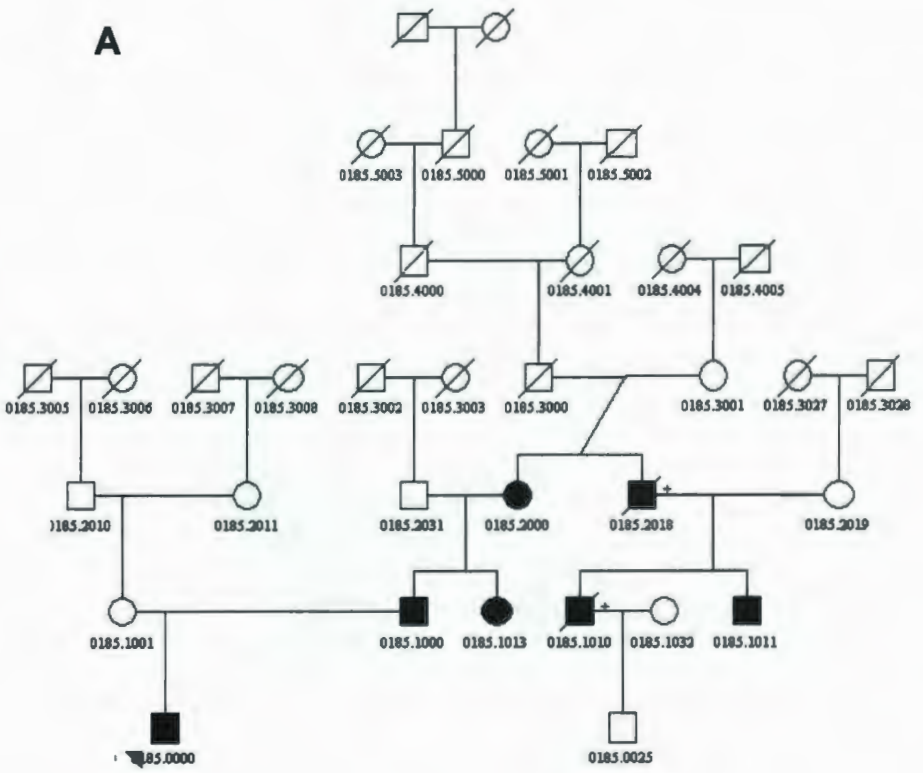
B2 – Family 64

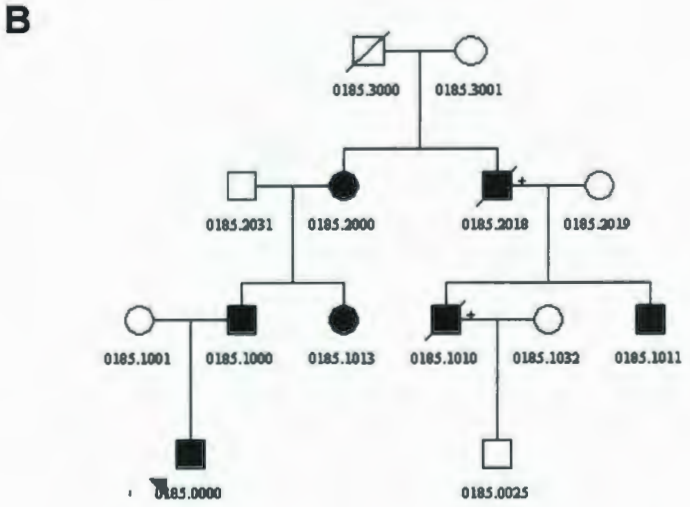

A
B
0069.2000
0069.2003
0069.1012
0069.1000
0069.1007
0069.1006
0069.0029
0069.0000
0069.0006
0069.0002
0069.0004
0069.0005
0069.0063
0069.0012
0069.0011
0069.0072
0069.A002
0069.A000
0069.A004
0069.A005
0069.A007
0069.A009
0069.A010
0069.A012
0069.A018
0069.A019
0069.A025
0069.A024
0069.A026
0069.A028
0069.A030
0069.B002
0069.B013
0069.B014
0069.B015

B3 – Family 69

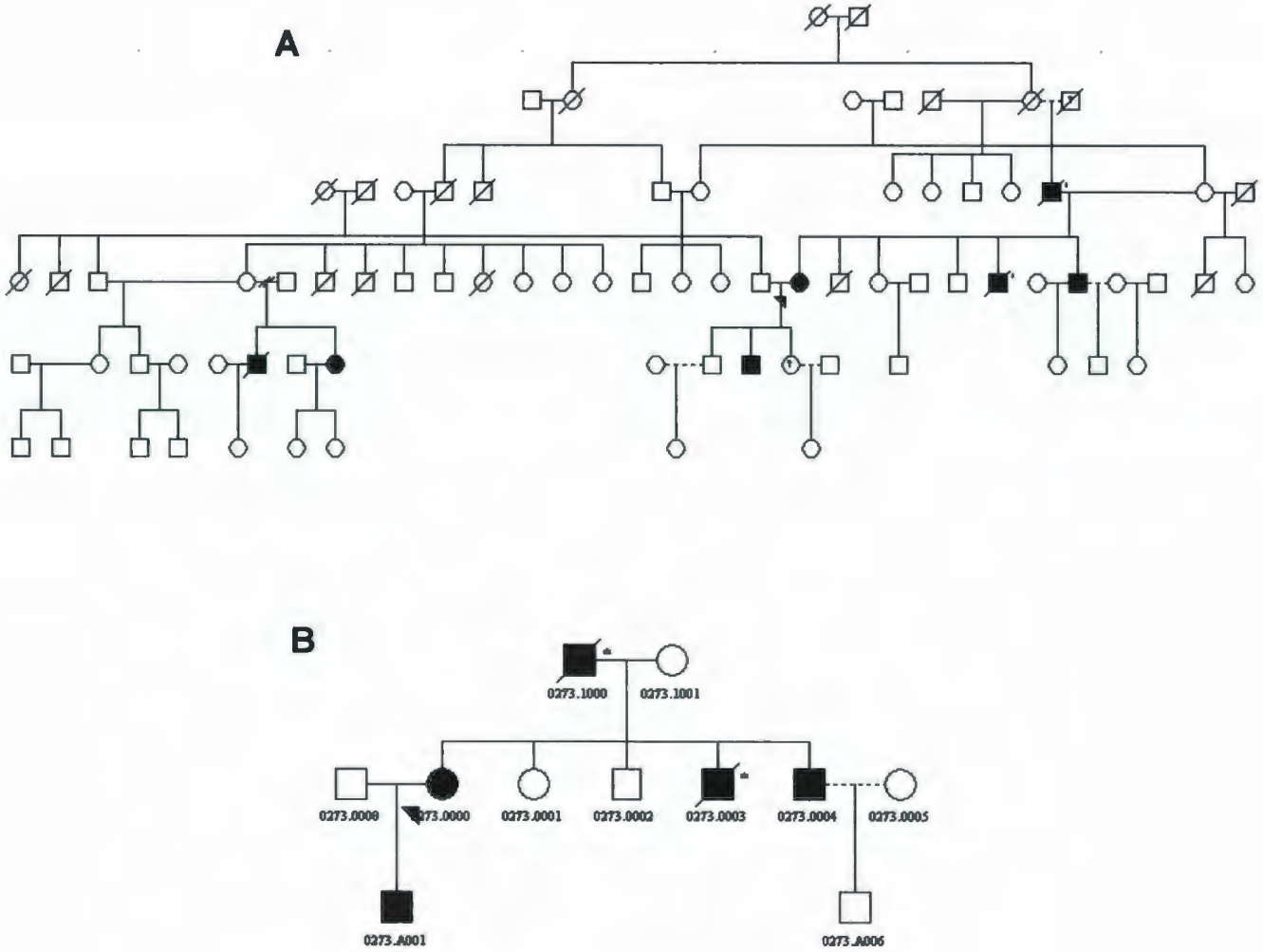

A
B
0076.3001
0076.3000
0076.3003
0076.3002
0076.3010
0076.3009
0076.2093
0076.2094
0076.2000
0076.2006
0076.2023
0076.2022
0076.2001
0076.2002
0076.1132
0076.1001
0076.1000
0076.1002
0076.1005
0076.1004
0076.0036
0076.1023
0076.1011
0076.1006
0076.1024
0076.1026
0076.1008
0076.1029
0076.1010
0076.1014
0076.0000
0076.0003
0076.0080
0076.0005
0076.0007
0076.0008
0076.0039
0076.0041
0076.0042
0076.0043
0076.0044
0076.0046
0076.0047
0076.0009
0076.0049
0076.0060
0076.0063
0076.0062
0076.0083
0076.0064
0076.0066
0076.A036
0076.A024
0076.A030
0076.A032
0076.A034
0076.A003
0076.A004
0076.A005

B4 – Family 76

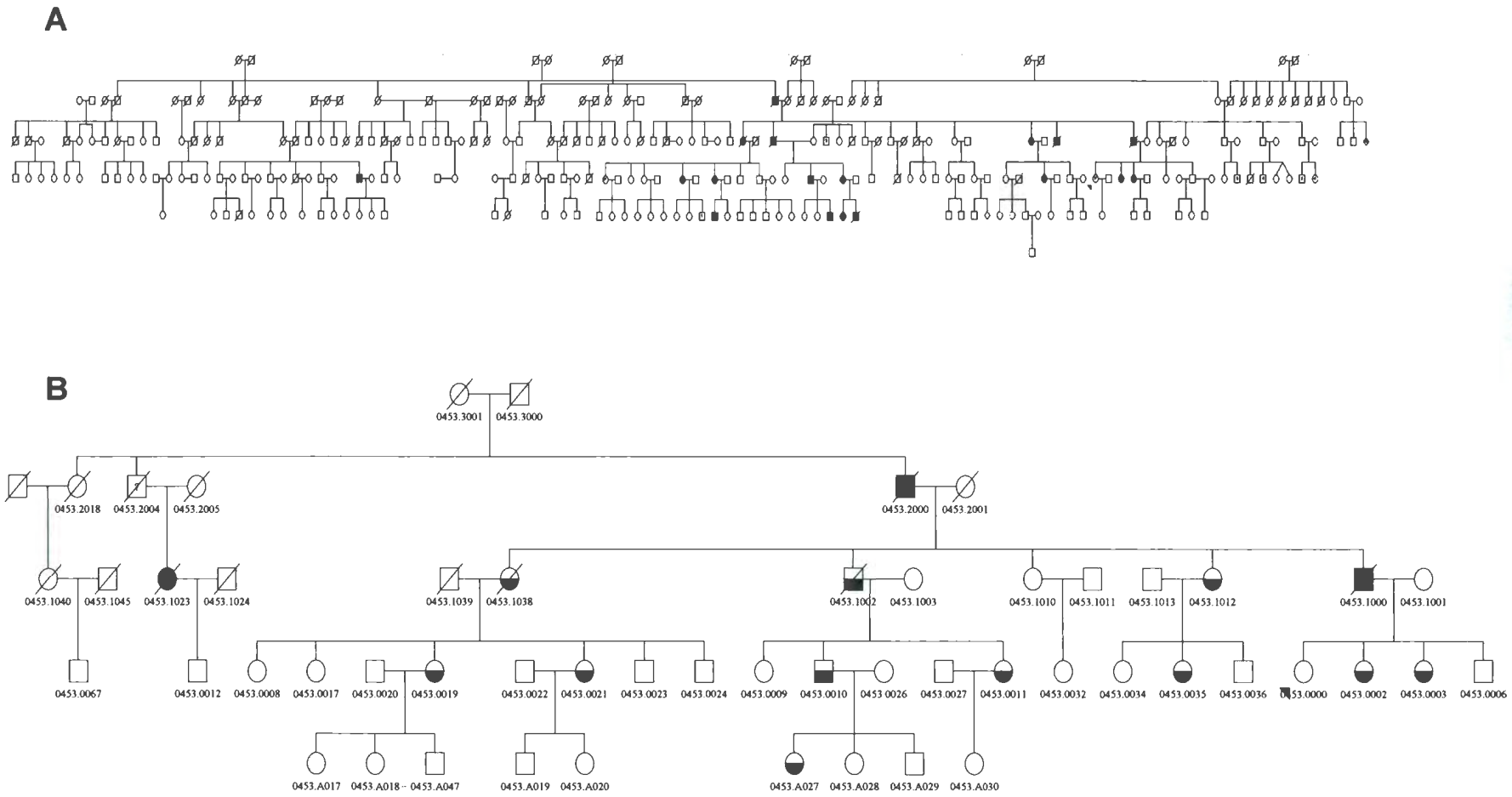




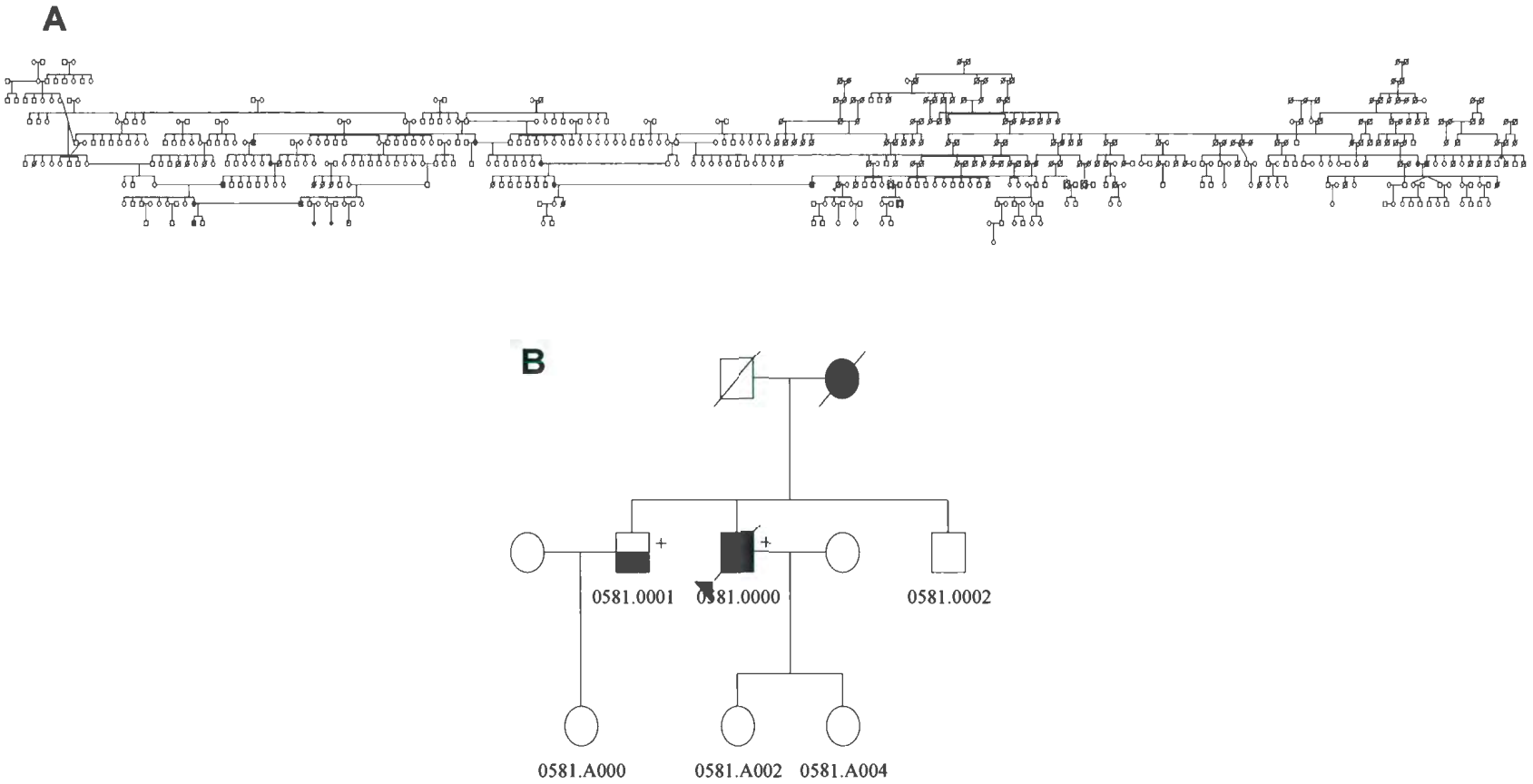
B5 – Family 185



B6 – Family 273

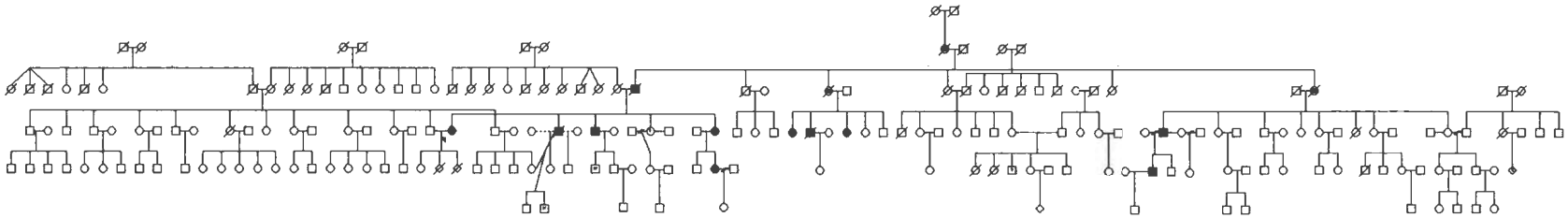


B7 – Family 453

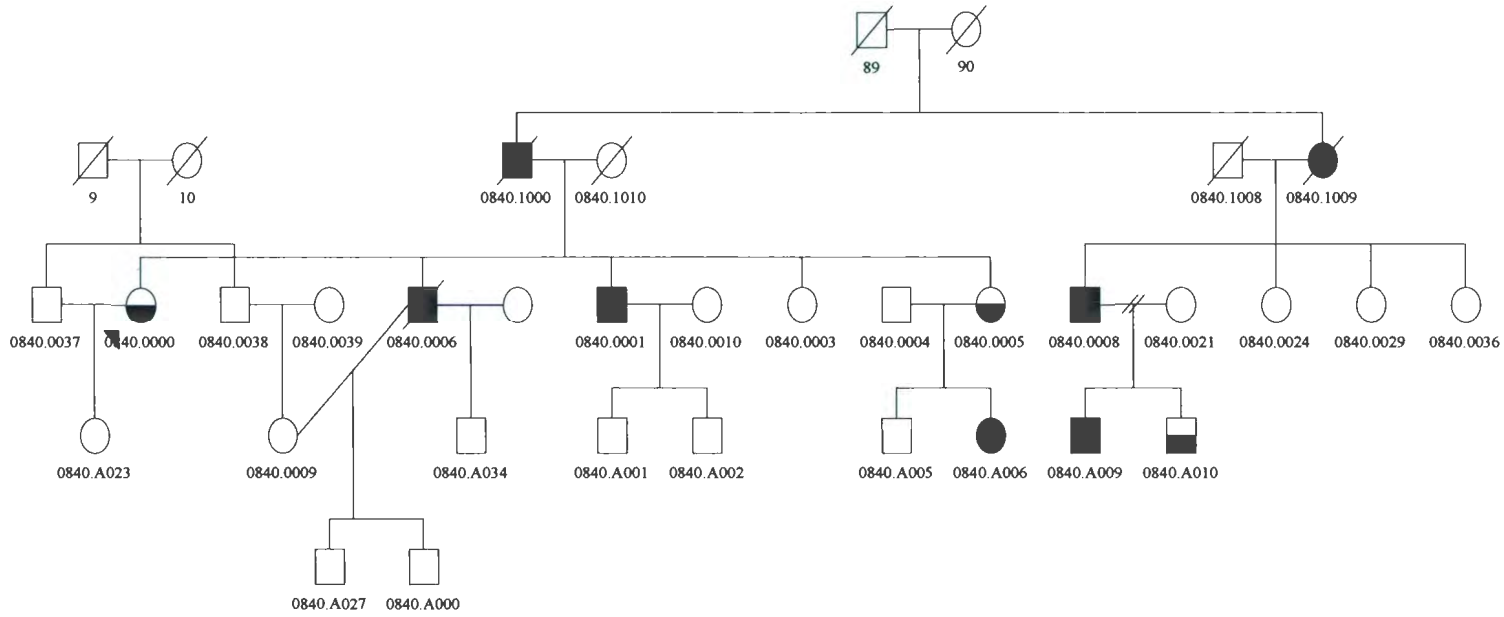


B8 – Family 581

A

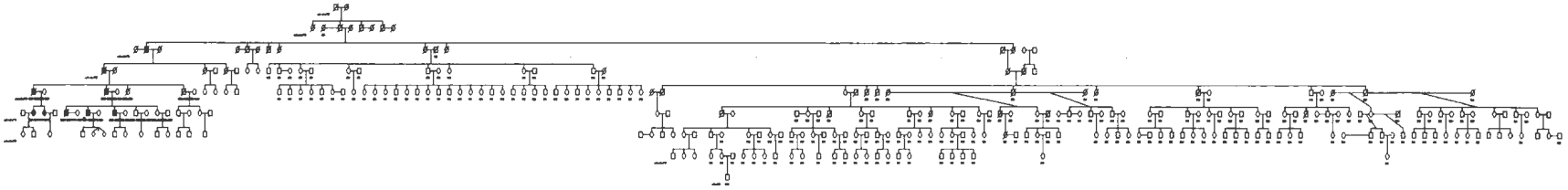


B

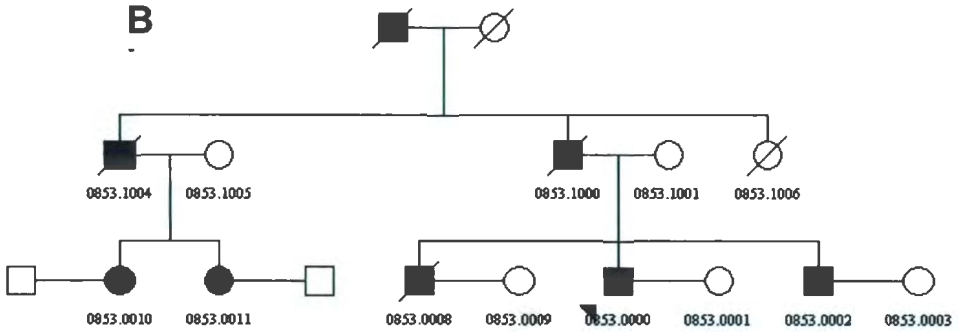


B9 – Family 840

A

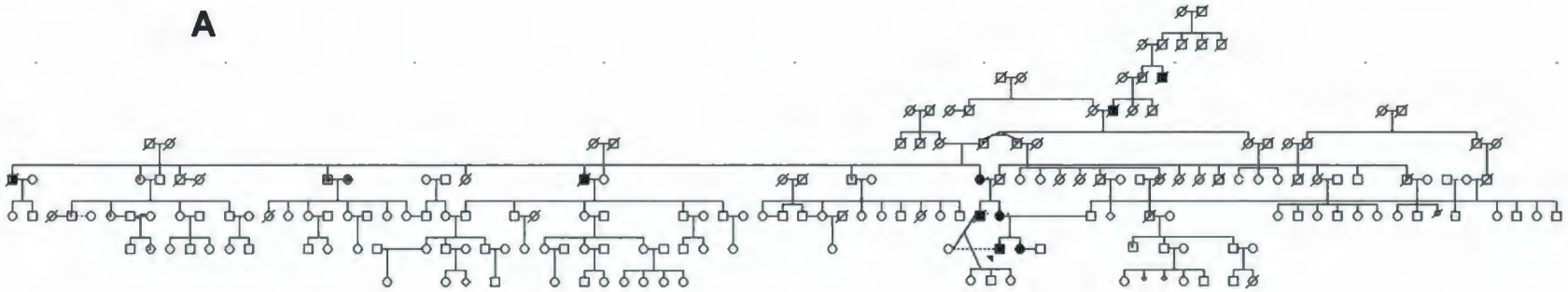


B

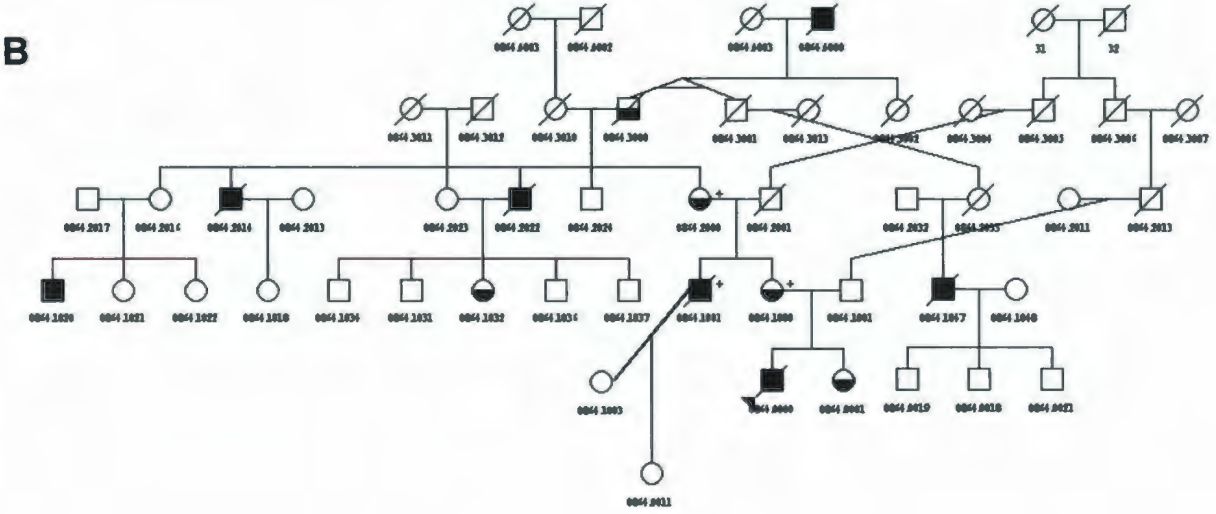


B10 – Family 853

A

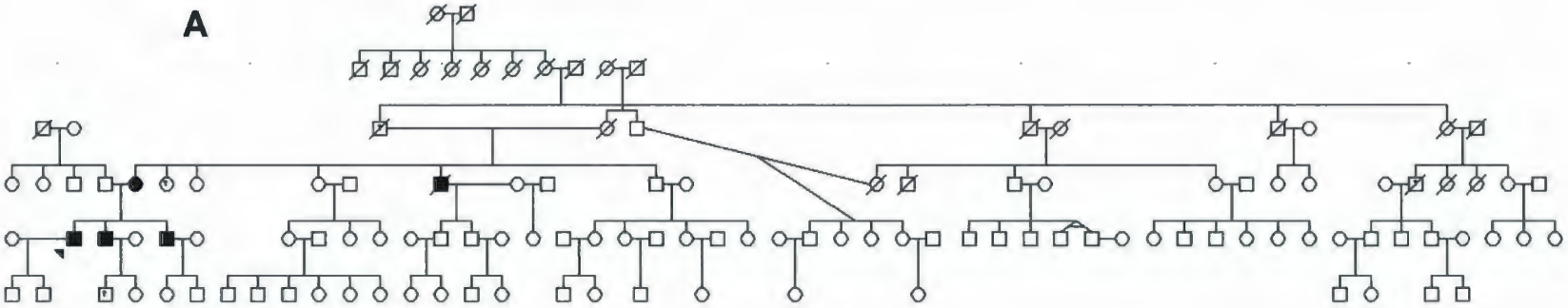


B

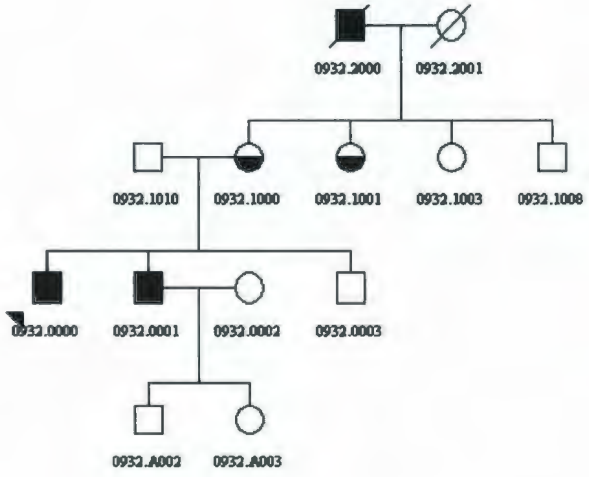


B11 – Family 864

A

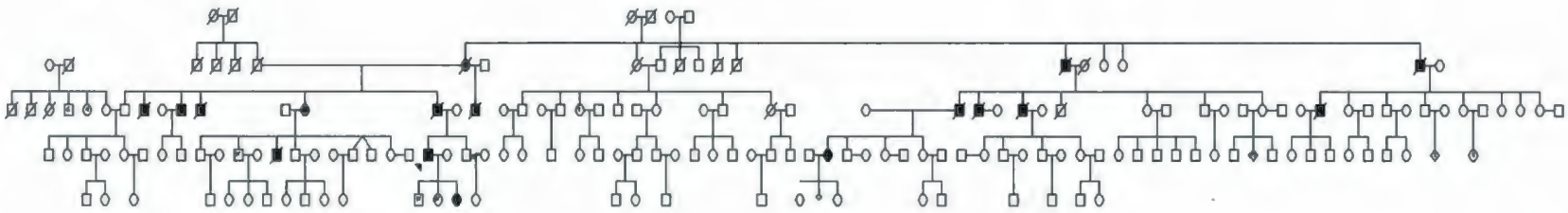


B

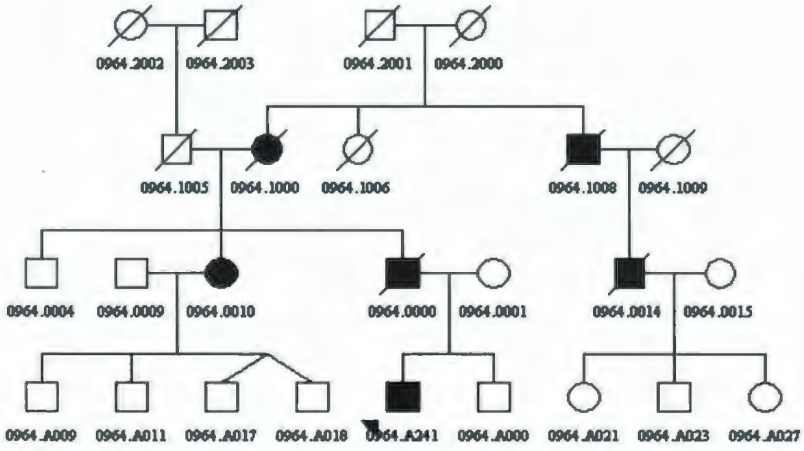


B12 – Family 932

A

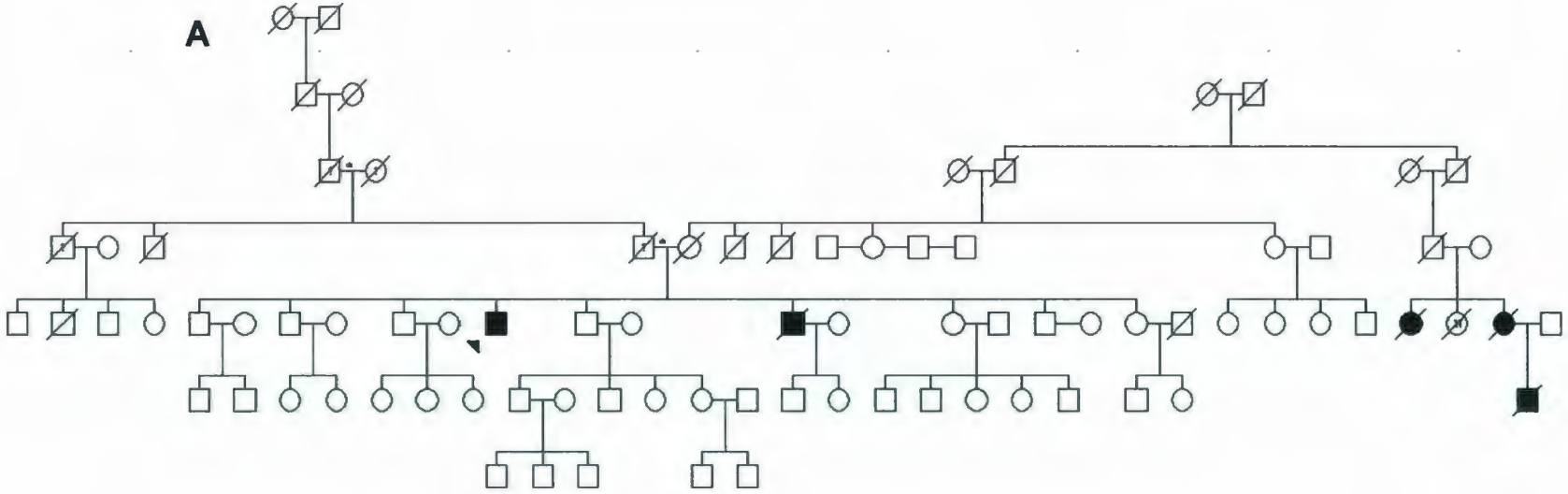


B

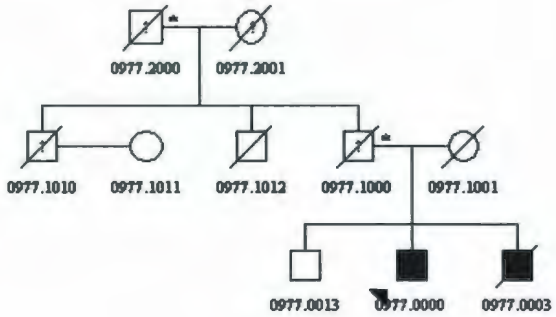


B13 – Family 964

A

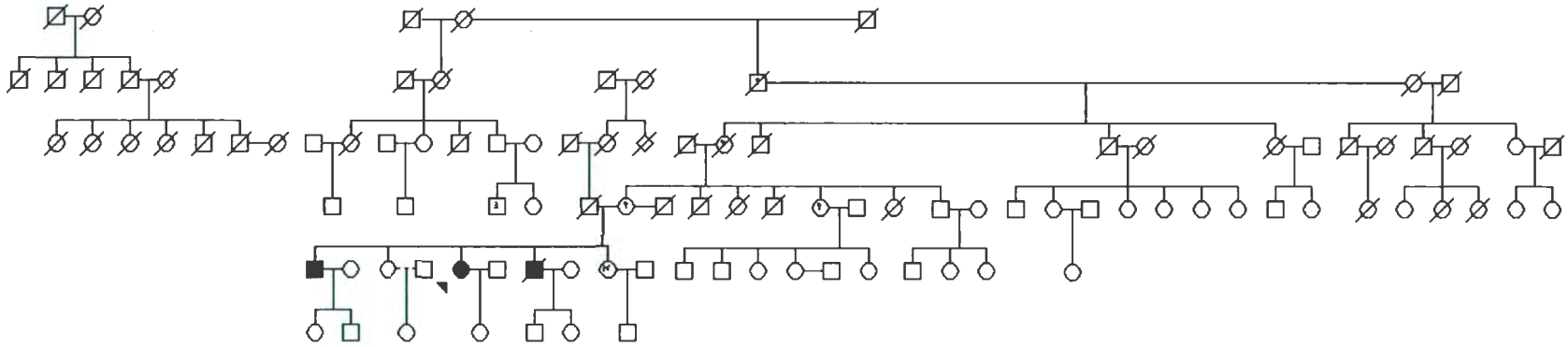


B

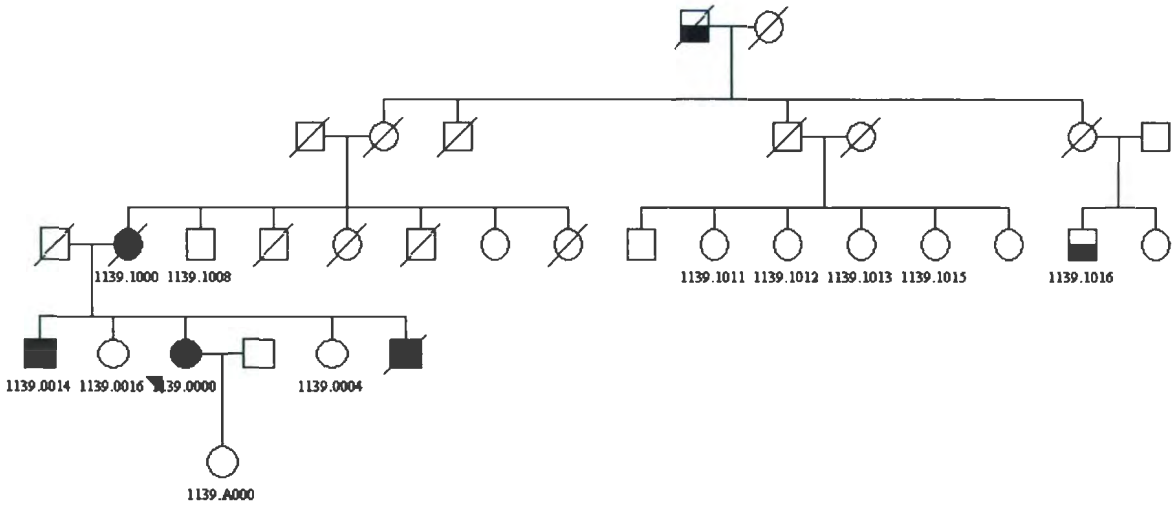


B14 – Family 977

A



B



B15 – Family 1139

Appendix C: Primers and PCR conditions for mutation screening

Gene	Exon	Primer Name	Primer Sequence	Amplicon Size (bp)	Anneal Temp (°C)	Betaine	MgCl (mM)
IQSEC1	1	NM_014869-Ex1F NM_014869-Ex1R	acttcagggccacattcctc CTCACCCAAACCCACAGTCT	518	TD58	Yes	1.5
	2	NM_014869-Ex2F NM_014869-Ex2R	ccatggggtgtgagaagaac agctacaggtggagctggaa	588	TD54	Yes	1.5
	3	NM_014869-Ex3aF NM_014869-Ex3aR	tgaccaactttgggaccttc CGTGTCAGCCTTGTCCTCTT	827	TD56	Yes	1.5
	3	NM_014869-Ex3bF NM_014869-Ex3bR	CATCGATGAGGAGGAGCTGT ttgtgtggcagcttcctatg	828	TD56	Yes	2.5
	4	NM_014869-Ex4#2F NM_014869-Ex4#2R	cagcctgggttcttgagttc acatgtcccagcaagtaggg	465	TD54	Yes	1.5
	5	NM_014869-Ex5F NM_014869-Ex5R	ctccccagaagactggaggt atgcagtctttggtccatcc	400	TD54	Yes	1.5
	6	NM_014869-Ex6F NM_014869-Ex6R	gagtgaggagggcaagtgag gtcatctgtgctccccaaag	500	TD54	Yes	1.5
	7	NM_014869-Ex7F NM_014869-Ex7R	tgggtctagtggaaggttcg tggaagagaccacacggagt	345	TD54	No	1.5
	8	NM_014869-Ex8F NM_014869-Ex8R	tcccaagggtaggtgtaggg aactcaagcagccagcactt	243	TD54	Yes	1.5
	9	NM_014869-Ex9F NM_014869-Ex9R	ctgtgagggaaggcattttt tgtgaaagacaccaggcaaa	400	TD54	Yes	1.5
	10	NM_014869-Ex10F NM_014869-Ex10R	tctgaccagaggtcccacat tgaggctcagctgactgatg	370	TD54	Yes	1.5
	11	NM_014869-Ex11F NM_014869-Ex11R	gcacaactcagccctaccat agggaaggccagaaaagact	388	TD54	Yes	1.5
	12	NM_014869-Ex12F	ggacagtgctagctccaagc	475	TD54	Yes	1.5

		NM_014869-Ex12R	tttgacgtttgggtcctgtt				
	13	NM_014869-Ex13F NM_014869-Ex13R	cctcatgtccgcgtctaact ctctgtgcagtaggggaagg	250	TD54	Yes	1.5
	14	NM_014869-Ex14aF NM_014869-Ex14aR	ctgtgctcttgtcgcttctg GGTCCCATGCTGTTTCTTTC	572	TD54	Yes	1.5
		NM_014869-Ex14bF NM_014869-Ex14bR	TCCAGAGAGAGCAAATACAG CA GTTCCAGCAGTTCCTTACGC	679	TD54	Yes	1.5
		NM_014869-Ex14cF NM_014869-Ex14cR	GAAAAGATTCAGGCCGTTTG tgcagtgtgtgagggaaaga	826	TD56	Yes	1.5
NUP210	1	NM_024923-Ex1-F NM_024923-Ex1-R	caagccttctccctcagctc CGCCATGACATGAGCAGT	624	TD56	Yes	1.5
	2	NM_024923-Ex2-F NM_024923-Ex2-R	cctcagcttgtggaggtctt tcctgccactcaacaagaga	467	TD54	Yes	1.5
	3	NM_024923-Ex3-F NM_024923-Ex3-R	cttggcttctgagggtgaag ctgaccagagctgccctcta	393	TD54	Yes	1.5
	4	NM_024923-Ex4-F NM_024923-Ex4-R	tggtgatgtggagtgtggtt ctggtgaggagctcacaggt	281	TD54	Yes	1.5
	5	NM_024923-Ex5-F NM_024923-Ex5-R	ggacacccctttatttcagca acgggacagcatctcacttc	397	TD54	Yes	1.5
	6	NM_024923-Ex6-F NM_024923-Ex6-R	gaaagtgaggatttgggaatga tgctacctgatggagcactt	398	TD54	Yes	1.5
	7	NM_024923-Ex7-F NM_024923-Ex7-R	gcgtcagctgtactgagagc gttgaggggaaacggctatt	394	TD54	Yes	1.5
	8	NM_024923-Ex8-F NM_024923-Ex8-R	cctaagccaggggagttttc tggtcatagcaagtggcaga	339	TD54	Yes	1.5
	9	NM_024923-Ex9-F NM_024923-Ex9-R	cttctcgtgtggatgctgag agcactgtttcccacagacc	375	TD54	Yes	1.5
	10	NM_024923-Ex10-F NM_024923-Ex10-R	tagcttcgtgcccaaattct atgcaagtagcccaaacagg	353	TD54	Yes	1.5

11	NM_024923-Ex11-F NM_024923-Ex11-R	ggcaccttcctctccactct gcaccctgccttattctcaa	500	TD54	Yes	1.5
12	NM_024923-Ex12-F NM_024923-Ex12-R	gagtgtgtgcatgggtgaac agtgaagtcgaggcttgctc	388	TD54	Yes	1.5
13	NM_024923-Ex13-F NM_024923-Ex13-R	gcacacagcagcacttagga aagccctctctttgggaatc	384	TD54	Yes	1.5
14	NM_024923-Ex14-F NM_024923-Ex14-R	aaagtgatttcagggtgacca aatttgcatctgggaaggag	248	TD54	Yes	1.5
15	NM_024923-Ex15-F NM_024923-Ex15-R	gctgcagatgacgagaatga aggagtggctgtctctcagc	487	TD54	Yes	1.5
16	NM_024923-Ex16-F NM_024923-Ex16-R	aaaactgagtcagagggagctg gaagtgcagcgtggagctat	390	TD54	Yes	1.5
17	NM_024923-Ex17-F NM_024923-Ex17-R	tttctggcatatcccattcc actaagctcccacaggcact	422	TD54	Yes	1.5
18	NM_024923-Ex18-F NM_024923-Ex18-R	cgaaactgcaaatcagaagga gactcaggacactgtgataccc	292	TD54	Yes	1.5
19&20	NM_024923-Ex19&20-F NM_024923-Ex19&20-R	agacctcaaagggtctgtgg tccgagtgtgagtcaagtgc	572	TD54	Yes	1.5
21	NM_024923-Ex21-F NM_024923-Ex21-R	gcactgctcacaaacctgag agggactctggatcctccac	397	TD54	Yes	1.5
22&23	NM_024923-Ex22&23-F NM_024923-Ex22&23-R	tgggtgactctcctgattcc tctctacagtagcttcctggactg	600	TD54	Yes	1.5
24&25	NM_024923-Ex24&25-F NM_024923-Ex24&25-R	gatgtccatgtcccaggttt gggtggtcctcagaggtaga	642	TD54	Yes	1.5

26	NM_024923-Ex26-F NM_024923-Ex26-R	atggtgctggtcatcctga gaagcaggccctagataccc	340	TD54	Yes	1.5
27	NM_024923-Ex27-F NM_024923-Ex27-R	aacctgtgccatgctcttct gcacctactcaccccgtcta	473	TD54	Yes	1.5
28	NM_024923-Ex28-F NM_024923-Ex28-R	cagctctggggttagcagac atcacacccggatgtttacc	500	TD54	Yes	1.5
29	NM_024923-Ex29-F NM_024923-Ex29-R	aggtcagcccttcacatcat gcacagcaggagaaaacctc	374	TD54	No	1.5
30	NM_024923-Ex30-F NM_024923-Ex30-R	gcatgaggagccacagactc atcttcggctctgatgacca	465	TD54	Yes	1.5
31	NM_024923-Ex31-F NM_024923-Ex31-R	gcattgttcctgaggtggat atgcccacatctcttggaaa	429	TD54	No	1.5
32	NM_024923-Ex32-F NM_024923-Ex32-R	gctcagcagaggacatggat cttgcagaaaacaggcacag	500	TD54	No	1.5
33	NM_024923-Ex33-F NM_024923-Ex33-R	gtctcctgaggctgctgagt tcagacggcttttccctaaa	383	TD54	No	1.5
34	NM_024923-Ex34-F NM_024923-Ex34-R	caaggcgtttaggctttgag ccctctctggagttggtgtg	393	TD54	No	1.5
35	NM_024923-Ex35-F NM_024923-Ex35-R	ctggtggaatttggtgtggt accctctgtgaccagcagtc	391	TD54	Yes	1.5
36	NM_024923-Ex36-F NM_024923-Ex36-R	gacttctcacctgggctttg taaagagagctggcccagga	477	TD54	Yes	1.5
37	NM_024923-Ex37-F NM_024923-Ex37-R	cagagatggccatgcagag ctgatagccaggaccactcc	399	TD54	No	1.5
38&39	NM_024923-Ex38&39-F NM_024923-Ex38&39-R	cctctgagaccaccatcctt gggcctgttgagacttaggg	489	TD54	No	1.5
40	NM_024923-Ex40a-F NM_024923-Ex40a-R	tgagagtgtcctgggtgagg CAGCACATTGTCCAGAAACC	499	TD54	Yes	1.5

		NM_024923-Ex40b-F NM_024923-Ex40b-R	TGCCGTCTCTTCACACAGAG GCCTAAGCCAAGTCCATCAG	599	TD54	Yes	1.5
		NM_024923-Ex40c-F NM_024923-Ex40c-R	GAGGCTGGAGAACACAGGA G CAGAATCAGGAGCTGGGAAG	538	TD54	No	1.5
		NM_024923-Ex40d-F NM_024923-Ex40d-R	CACTTTCATGGCAGCTCATC ccaggccccagagtatcat	567	TD54	No	1.5
HDAC11	1	NM_024827-Ex1-F NM_024827-Ex1-R	ccagaggacgcggctaag acccctcaccagcaaatct	240	TD54	Yes	1.5
	2	NM_024827-Ex2-F NM_024827-Ex2-R	aaaggcagccagtttaagca cgagggctgaaactgaagat	426	TD54	Yes	1.5
	3	NM_024827-Ex3-F NM_024827-Ex3-R	aagccagcagtgcctaccta ccacagactcccatgttcct	386	TD54	Yes	1.5
	4	NM_024827-Ex4-F NM_024827-Ex4-R	gtgtgcttgaccagggagtt ggagaaagggtcattgttgc	400	TD54	Yes	1.5
	5	NM_024827-Ex5-F NM_024827-Ex5-R	ccagagctggctagaagcag ggtgacttgggaaaaccaga	284	TD54	Yes	1.5
	6	NM_024827-Ex6-F NM_024827-Ex6-R	gaggcacaggttcctgagag cagcataggccccttgtcta	308	TD54	Yes	1.5
	7	NM_024827-Ex7-F NM_024827-Ex7-R	gattacccacaagcattaggc aagacagaagccgtgagagc	298	TD54	Yes	1.5
	8	NM_024827-Ex8-F NM_024827-Ex8-R	caaatggggagtttcctgag aggatcacctcacagggatg	320	TD54	Yes	1.5
	9	NM_024827-Ex9-F NM_024827-Ex9-R	gctgagcagtgctgaatctg cagtccatctcctccctgac	393	TD54	Yes	1.5
	10	NM_024827-Ex10a-F NM_024827-Ex10a-R	agcaggacttcctgacacca GGACCTAGCCTGTCCTCTCC	584	TD54	Yes	1.5
		NM_024827-Ex10b-F NM_024827-Ex10b-R	TTCTAACCTCATGGGGTGGT ttgcactgaacaggcaagac	683	TD54	Yes	1.5
Fibulin2	1	1-NM_001004019-F	gtcagcgagtctgggctct	463	TD50	Yes	1.5

	1-NM_001004019-R	GTGCCCGCATCTGTACACT				
2	2a-NM_001004019-F 2a-NM_001004019-R	tgaggcagtgcttagtgtgc AGCTGTCAGCCACCACTACA	577	TD50	Yes	2
	2b-NM_001004019-F 2b-NM_001004019-R	GCCGGTCAGTCCTATTTTGT TCCTCGGTCACTCTCCTAGC	575	TD50	Yes	2
	Ex2C#2-NM_001004019-F Ex2C#2-NM_001004019-R	GAGGTGGGAGTCAGCCACT GCAGTGATGTGGACAGGATG	484	TD54	Yes	1.75
3	3-NM_001004019-F 3-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	264	TD54	Yes	2
4	4-NM_001004019-F 4-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	324	TD54	Yes	1.5
5	5-NM_001004019-F 5-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	358	TD50	Yes	2.5
6	6-NM_001004019-F 6-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	399	TD50	Yes	2
7	7-NM_001004019-F 7-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	300	TD50	Yes	2
8	8-NM_001004019-F 8-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	296	TD54	Yes	2
9	9-NM_001004019-F 9-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	294	50	Yes	1.5
10	10-NM_001004019-F 10-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	397	TD56	No	1.5
11	Ex11#2-NM_001004019-F Ex11#2-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	381	TD54	Yes	1.5
12&13	12&13-	caaaatcctctcccaaatgc	590	TD50	Yes	2

	NM_001004019-F 12&13- NM_001004019-R	ggcatgaagcacattcacac				
14	Ex14#2- NM_001004019-F Ex14#2- NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	374	TD54	Yes	2.5
15	15-NM_001004019-F 15-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	299	TD56	No	1.5
16	16-NM_001004019-F 16-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	284	TD50	Yes	2
17	Ex17#2- NM_001004019-F Ex17#2- NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	482	TD54	Yes	1.5
18	Ex18a#2- NM_001004019-F Ex18a#2- NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	691	TD54	Yes	1.5
	18b-NM_001004019-F 18b-NM_001004019-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	647	TD56	Yes	4.5
EST1	BC111426-F BC111426-R	caaaatcctctcccaaatgc ggcatgaagcacattcacac	787	TD56	Yes	1.5
EST2	W70042_FBLN-est2-F W70042_FBLN-est2-R	ATATTTCCTGCCGAGCTGAG GCCAGGTCTTCCTCACACAT	599	TD54	Yes	1.5
EST4	DA856838-FBLN-Est4- F DA856838-FBLN-Est4- F	GGGAGGGAGTGGAGACAGA T ctgtaagaccctgggggaat	815	TD54	Yes	1.5
EST5	DB458158-FBLN-est5- ex1-F	ctggagggagtgttctccac	247	TD54	Yes	1.5

		DB458158-FBLN-est5-ex1-R	tggtgctgtctgatggttgt				
	EST6	BF368851-FBLN2-Est6-F BF368851-FBLN2-Est6-R	ccagccgagactgttctgat ctggaccccagctttcact	342	TD54	Yes	1.5
WNT7A	1	NM_004625-Ex1a#2 NM_004625-Ex1a#2	gccctgcaatctgcaagtta gcggggcaatcaacatag		TD57 (40 cycles)	Yes	1.5
		NM_004625-Ex1b NM_004625-Ex1b	cgtctcgcacacttgcac tccccctgcctatatctcct	527	TD54	Yes	1.5
	2	2-NM_004625-F 2-NM_004625-R		384	TD56	Yes	2
	3	3-NM_004625-F 3-NM_004625-R		498	TD56	Yes	2
	4	NM_004625-Ex4a#1-F NM_004625-Ex4a#1-R	aagccagacccagataggtg ggtcctcctcgcagtagttg	387	TD54	Yes	1.5
		NM_004625-Ex4a#2-F NM_004625-Ex4a#2-R	gaccacactgccacagtttc tcctcccagcaatctgactt	455	TD54	Yes	1.5
		Exon4b-NM_004625-F Exon4b-NM_004625-R	gcggctgtgacctcatgt tctacacggctccttgtcct	593	TD56	Yes	2
	EST1 exon 1	AA405144-1-F AA405144-1-R	cctctgtgagcctcaattttc tgttttgtttactggatattcttga	399	TD56	Yes	3
	EST1 exon 2	AA405144-2-F AA405144-2-R	ccatcccagcatcttcctt ctggcctagaggatgtgtgc	246	TD54	No	1.5
	EST2 exon 2	AA412464-2-F AA412464-2-R	ctgggcttctgcctctgtag tgaactgaggcttgcaaaaa	399	TD54	No	1.5

	EST3	AA826310WNT-est3-F AA826310WNT-est3-R	cccttcatgtttgagggtgt tggatgcactgacttcaAGC	686	TD54	Yes	2.5
TPRXL	1	AK092426-Ex1-F AK092426-Ex1-R	ctggtgtgttcaggaagctg gtaaagccactcaccccaag	279	TD54	Yes	1.5
	2	AK092426-Ex2-F AK092426-Ex2-R	ttcctcagcctttgtttttg cccgccagacatacatatcc	497	TD56	Yes	2.5
	3	AK092426-Ex3-ext-F AK092426-Ex3-ext-R	ggcatcttttccctttgtaatg gaagctggaaggcacagttc	2173	TD54 ext 1:00	Yes	1.5
		AK092426-Ex3-int-F AK092426-Ex3-int-R	gtgcagagccaggagggcgtttg cctaggattgagcctggacctg				
CHCHD4	1	BC033775-Ex1-F BC033775-Ex1-R	gaaatgtagtttccggctgag ATCCGCAAAACAGGAGTAGC	397	TD54	Yes	1.5
	2	BC033775-Ex2-F BC033775-Ex2-R	ttcaaacttggtttgtgtgg ccagggaagtacatcagaacaa	400	TD54	Yes	1.5
	3	BC033775-Ex3-F BC033775-Ex3-R	gctaaatgggtcccggtatt ctgtgtgaaggcatctgctc	1595	TD58	Yes	1.5
		BC033775-Ex3-internal-F BC033775-Ex3-internal-R	ggccttgtcatcaccttccaag tcagggcaatctgagaattc				
	2	NM_144636-Ex2-F NM_144636-Ex2-R	tcctgaggcttcagacatcc caactcccccatacagaatca	380	TD54	Yes	1.5
TMEM43	1	NM_024334-Ex1-F NM_024334-Ex1-R	caatgtcccggaccgtatag ggcgaaatggacctagagga	320	TD54	Yes	1.5
	2	NM_024334-Ex2-F NM_024334-Ex2-R	agttttcattctgttactgtttctttt ggcccttgattaccaaatcc	282	TD54	Yes	1.5
	3	NM_024334-Ex3-F NM_024334-Ex3-R	aactgtacggtggggagatg atcactcccatgtgtgacca	375	TD54	Yes	1.5
	4	NM_024334-Ex4-F NM_024334-Ex4-R	aagaacctgggacagggagt ctcctggagccactcttcac	472	TD54	Yes	1.5

	5&6	NM_024334-Ex5&6-F NM_024334-Ex5&6-R	tgatctggtagccctgaggt cacgaggcaggattaactcaa	596	TD54	Yes	1.5
	7	NM_024334-Ex7-F NM_024334-Ex7-R	cctgggctaatctggacttg ctgatcctgtgcctttagcc	300	TD54	Yes	1.5
	8&9	NM_024334-Ex8&9-F NM_024334-Ex8&9-R	cgtggacgagacagagtcag cgctcctgacattgaccaag	700	TD54	Yes	1.5
	10	NM_024334-Ex10-F NM_024334-Ex10-R	gggtttctgtgctcacttcc tgcctcattcactggctatg	330	TD54	Yes	1.5
	11	NM_024334-Ex11-F NM_024334-Ex11-R	tgttcagaaatggccaacag ctcatcccaaggctatggag	394	TD54	Yes	1.5
	12	NM_024334-Ex12a-F NM_024334-Ex12a-R	cccatcctcatctagggaca GGAAACAGCAGGAGAAGCT G	559	TD54	Yes	1.5
		NM_024334-Ex12b-F NM_024334-Ex12b-R	TGGTGTTCACCAGCTCATGT TTCCTCTTGGGGTAGGAAAG	581	TD54	Yes	1.5
		NM_024334-Ex12c-F NM_024334-Ex12c-R	TCCTGAGGAGAAAAGCTGGA CGTGGGCATTGTACAACCAG	581	TD54	Yes	1.5
		NM_024334-Ex12d-F NM_024334-Ex12d-R	CGATTAAGAGAAAAGGTTGG AA GAGATTTGATGAAATTGCTC ATGTA	592	TD54	Yes	1.5
		NM_024334-Ex12e-F NM_024334-Ex12e-R	TTGTGCCTGCTGGGAGTAAT atcctatggctgaattctttaca	561	TD54	Yes	1.5
XPC	1	NM_004628-Ex1-F NM_004628-Ex1-R	ttcctttaggggcgtgacta agctacgcaggagcttggat	457	TD54	No	1.5
	2	NM_004628-Ex2-F NM_004628-Ex2-R	ggggtttggagacaggtcat tggaccccagtgacaagtaag	399	TD54	No	1.5
	3	NM_004628-Ex3-F NM_004628-Ex3-R	tggaggaagtgaggctcaga tgcaattagtgatctgactccaa	398	TD54	No	1.5
	4	NM_004628-Ex4-F	ttcccagcagaaccttgatt	300	TD54	No	1.5

	NM_004628-Ex4-R	tcccctacaagtttctccaaa				
5	NM_004628-Ex5-F NM_004628-Ex5-R	gccttgtgtagggaaacagg agagcagcaaagccagaaat	297	TD54	No	1.5
6	NM_004628-Ex6-F NM_004628-Ex6-R	tttcttatatgtagaaatggcaacac tagttaccgcctcagggaag	390	TD54	No	1.5
7	NM_004628-Ex7-F NM_004628-Ex7-R	aaagtctgagctgggtctgc tgtcggtaacacacctggaa	391	TD54	No	1.5
8	NM_004628-Ex8-F NM_004628-Ex8-R	cctgtcttgaacaagcaccat acccacactccgtgaatacc	296	TD54	No	1.5
9	NM_004628-Ex9a-F NM_004628-Ex9a-R	ggggacatcttgatgtattgg GGGTCCTGGAGGCACTCT	595	TD54	No	1.5
	NM_004628-Ex9b-F NM_004628-Ex9b-R	CCTCTGATGAGGATTCCGAA C gctgggcatatataaggtgctc	595	TD54	No	1.5
10	NM_004628-Ex10-F NM_004628-Ex10-R	gctccaccatctgttgtcag tgctgtccagtcagatgagc	375	TD54	No	1.5
11	NM_004628-Ex11-F NM_004628-Ex11-R	gcctagcacagcttctctgg cctccttgaatcctgctcaa	397	TD54	No	1.5
12&13	NM_004628-Ex12&13-F NM_004628-Ex12&13-R	ttctgagggttcaccaggta aggcctcaactcccagcag	600	TD54	No	1.5
14	NM_004628-Ex14-F NM_004628-Ex14-R	tcaggacccctagtgctcag agcctgctgtattcagtgctc	368	TD54	No	1.5
15	NM_004628-Ex15-F NM_004628-Ex15-R	tctgtctttacattcacagtttcca cacaaagctatccctgacttga	294	TD54	No	1.5
16	NM_004628-Ex16a-F NM_004628-Ex16a-R	ggaacttgctgcctcttcat ATGCACCACCATCCAGAAAT	683	TD54	Yes	1.5
	NM_004628-Ex16b-F NM_004628-Ex16b-R	CAGCCCTTGTCAGATTCACC ccctcaggaagaccactcaa	699	TD54	Yes	1.5

LSM3	1	1-NM_014463-F 1-NM_014463-R	ggaagcaggaaagagcacaa ggccatttttcctgagtctg	221	TD54	No	1.5
	2	2-NM_014463-F 2-NM_014463-R	cccatggtcacacagcatag acccttcactctggacagg	280	TD54	No	1.5
	3	3-NM_014463-F 3-NM_014463-R	aaatgccctagttgtaattgagc tcaaaacatgacaggcttcc	292	TD54	No	1.5
	4	4-NM_014463-F 4-NM_014463-R	tttgtgtcatgaatccattt caaggaaacaccaatttaaagaaaa	490	TD56	Yes	2
SLC6A6	1	NM_003043-Ex1 NM_003043-Ex1	cgtgtttgcgtgtatgtgtg GCCGGGGAAACAATAACAG	589	TD54	Yes	1.5
	2	NM_003043-Ex2 NM_003043-Ex2	gctgggagtctgcatatgatt accaagcaggcacagagtg	480	TD54	Yes	1.5
	3	NM_003043-Ex3 NM_003043-Ex3	caaagcattggacagacagg ggaaagcagggaaagagagg	492	TD54	Yes	1.5
	4	NM_003043-Ex4 NM_003043-Ex4	gtccaagagtttggccttga cctggcagacaataggtgct	351	TD54	Yes	1.5
	5	NM_003043-Ex5-BC111489 NM_003043-Ex5-BC111489	ataccatggccttccagttg caagtctggggatggttgtt	793	TD54	Yes	1.5
	6	NM_003043-Ex6 NM_003043-Ex6	ctgttaatgtccgagccttg ctaattgcaggcgggaagt	295	TD54	Yes	1.5
	7	NM_003043-Ex7 NM_003043-Ex7	ctcccctctgcctctctga cctctacatcctgccctgtc	298	TD54	No	1.5
	8 & 9	NM_003043-Ex8&9 NM_003043-Ex8&9	TTGAAAGGCACCTTGGATTC TTCATGCCTGCAGTACTTGG	591	TD54	Yes	1.5
	10	NM_003043-Ex10 NM_003043-Ex10	ccctgctgtatttgggtctg cccaggtggaataaggagtct	300	TD54	Yes	1.5
	11	NM_003043-Ex11 NM_003043-Ex11	atctcacaggccaattctgg ggagatgaagcccctacctg	385	TD54	Yes	1.5

12	NM_003043-Ex12 NM_003043-Ex12	tgccagagctctttgtatatga agggagggcctactgttcaa	300	TD54	Yes	1.5
13	NM_003043-Ex13 NM_003043-Ex13	gtcccagagaccccactcat ttataagcccacgctgaaca	294	TD54	Yes	1.5
14	NM_003043-Ex14 NM_003043-Ex14	gagagacccttccctcaggt ctggatgggtgagggaacta	399	TD54	Yes	1.5
15	NM_003043-Ex15a NM_003043-Ex15a	acagcacatgagccaattca ATCTCCTCACAGCCCTCCTT	839	TD54	Yes	1.5
	NM_003043-Ex15b NM_003043-Ex15b	CCACTTGAATTGATCTTCTTG C TCCTTTGGTCCACTCACCTC	761	TD54	Yes	1.5
	NM_003043-Ex15c NM_003043-Ex15c	CATTCCAGGCAGAGAAGGAG TGTGAAAATTCTGCGGTCTG	810	TD54	Yes	1.5
	NM_003043-Ex15d NM_003043-Ex15d	GATCAAGGGCCTTATGTGGA GAGGGCTCTGATTGGAGACA	850	TD54	Yes	1.5
	NM_003043-Ex15e NM_003043-Ex15e	GCCACAGTATTTTGGGTTGG GGTCCAGGAATTCTGTGAGG	818	TD54	Yes	1.5
	NM_003043-Ex15f NM_003043-Ex15f	GCAGGGAATGGTGAGTGTCT CGATTTTTGCACACAGTGGT	843	TD54	Yes	1.5
	NM_003043-Ex15g NM_003043-Ex15g	TGCTCTGGGTAGCCAGTCTA A CCTGCAAAATACAGGGAATC A	761	TD54	Yes	1.5
	NM_003043-Ex15h NM_003043-Ex15h	GTCTTGCCTCCTTCCAGAAA gaggagcgtcctgggatt	456	TD54	Yes	1.5
mRNA1	BC038790-Ex1a BC038790-Ex1a	AGACTCATGCACAGCCTCCT GCTCAGACTTCAGAGCTGAC AA	581	TD54	Yes	1.5
	BC038790-Ex1b BC038790-Ex1b	GTCAAAGCACCTGGGAAATG TCACCTTGGGGGATAAAATG	816	TD54	Yes	1.5

		BC038790-Ex1c BC038790-Ex1c	CTTCGTGACCTGGAGGAGTC CAACAAGGTCAGCAAACTGG	828	TD54	Yes	1.5
	mRNA2	AK023516-Ex1a AK023516-Ex1a	CTAAGCAGGCCCTGAAAGC GCTTCAGTTCTCATGGATGG	700	TD54	Yes	1.5
		AK023516-Ex1b AK023516-Ex1b	CCACTGTAAACCACCGTCCT GCCTTCGTGGAACTGACATT	846	TD54	Yes	1.5
		AK023516-Ex1c AK023516-Ex1c	GGGATTTGAACCCTTGCTTA GAGTTCTGGAGCCAGGTTTG	750	TD54	Yes	1.5
GRIP2	1	NM_001080423-Ex1-F NM_001080423-Ex1-R	aaggggcagatctaactccaa agcagtgatgccaaggtctc	366	TD54	No	1.5
	2	NM_001080423-Ex2-F NM_001080423-Ex2-R	gagccagccttgacaatagc caggtctcagccatccagtc	327	TD54	No	1.5
	3	NM_001080423-Ex3-F NM_001080423-Ex3-R	atcttacaggccacgtccac gtagggcactctgctgcatt	374	TD54	No	1.5
	4	NM_001080423-Ex4-F NM_001080423-Ex4-R	gccactgacttgctctgtga acaccaagttgatgggtgct	471	TD54	No	1.5
	5	NM_001080423-Ex5-F NM_001080423-Ex5-R	tgaccctggacaactgtctg tactgcggttgtaggcacag	477	TD54	Yes	1.5
	6	NM_001080423-Ex6-F NM_001080423-Ex6-R	ctgatcacccagtcctcctg ttcaaagtcagaactgagatacaca	299	TD54	No	1.5
	7	NM_001080423-Ex7-F NM_001080423-Ex7-R	ggacaaggtcctcaaacagg gcctcgatctcttcatctgg	300	TD54	No	1.5
	8	NM_001080423-Ex8-F NM_001080423-Ex8-R	cattccctgcttcttcatcc gactggcccaaggtcataag	375	TD54	No	1.5
	9&10	NM_001080423-Ex9&10-F NM_001080423-Ex9&10-R	aaatgggtataacagtttcttcca tgctgctcttaataatttgagctt	700	TD54	No	1.5
	11	NM_001080423-Ex11-F NM_001080423-Ex11-	cgggacatgtgagcagaata actcactgcagcctcaacct	468	TD54	No	1.5

	R					
12	NM_001080423-Ex12-F NM_001080423-Ex12-R	aactgatgtgaaggggaag tggagctctgcccaaaatac	383	TD54	Yes	1.5
13	NM_001080423-Ex13-F NM_001080423-Ex13-R	ggaatcaacgatcccattgt cttgtgaagctagggcttgg	466	TD54	Yes	1.5
14	NM_001080423-Ex14-F NM_001080423-Ex14-R	gtgtgacccaggttcagtcc aaggcttcccttcccactat	400	TD54	No	1.5
15	NM_001080423-Ex15-F NM_001080423-Ex15-R	gcaggggggatgagattagaa atggaaatgctggtctgagc	299	TD54	No	1.5
16&17	NM_001080423-Ex16&17-F NM_001080423-Ex16&17-R	agtgggaccaggatgtgagt cccctccttcatttgttcat	584	TD54	No	1.5
18	NM_001080423-Ex18-F NM_001080423-Ex18-R	gaaatcagagcgtgagtgactg ccaatgtcacccacagtgtc	399	TD54	No	1.5
19	NM_001080423-Ex19-F NM_001080423-Ex19-R	gtttttggcagctgggttta ggctgcagtgaggactctgt	380	TD54	No	1.5
20	NM_001080423-Ex20-F NM_001080423-Ex20-R	gagtttccagaccaggctga taaaggatgccaggagagga	390	TD54	Yes	1.5
21	NM_001080423-Ex21-F NM_001080423-Ex21-R	gtagatgcaggcccagagag ggctctgctcagtgcttctc	450	TD54	Yes	1.5
22	NM_001080423-Ex22-F NM_001080423-Ex22-	gtctcatctgcccctgttct cacatgaatccatgcacagtc	391			

		R					
	23	NM_001080423-Ex23-F NM_001080423-Ex23-R	ggttcatgacttcccactcc ctctggagtcccacgacagt	394	TD54	Yes	1.5
	24	NM_001080423-Ex24-F NM_001080423-Ex24-R	taagggctccgtcgtacttg ttcacaggcgatagatgcag	469	TD54	No	1.5
	25	NM_001080423-Ex25-F NM_001080423-Ex25-R	acatcacatgcctcttgctc gtgtctgggagccaggatct	300	TD54	Yes	1.5
C3orf19	1	NM_016474-Ex1-F NM_016474-Ex1-R	tacgcaccacactctgacct gaaacaagctaggtcgacattc	299	TD54	Yes	1.5
	2	NM_016474-Ex2-F NM_016474-Ex2-R	ttgtccatttagcggaacttg catttcatgcagatgttaaaacaa	360	TD56	Yes	2.5
	3	NM_016474-Ex3+BF698254-Ex3-F NM_016474-Ex3+BF698254-Ex3-R	aagaactcatgatctcctttgttt cacaccatgggaaaacctct	483	TD54	Yes	1.5
	4	NM_016474-Ex4-F NM_016474-Ex4-R	ctgtcaggcagggagtgtg tgacgtctaaattttctctctactcct	248	TD54	Yes	2
	5	NM_016474-Ex5+CN278581-Ex5-F NM_016474-Ex5+CN278581-Ex5-R	caccaggcaatgaatgtgac caggcactgggtttccttc	552	TD54	Yes	1.5
	6	NM_016474-Ex6+CD58017981-Ex6-F NM_016474-Ex6+CD58017981-Ex6-R	ccccttcatcactgttcaccta cagtcaggttgttccccaag	386	TD54	Yes	1.5
	7	NM_016474-Ex7-F	tgggtcagaagcctctgtagt	396	TD54	Yes	1.5

		NM_016474-Ex7-R	caaagcaggatctttaatttcca				
	8	NM_016474-Ex8-F NM_016474-Ex8-R	aggaaatgtggcttgagtgt ggcaactacctttatgcaattc	300	TD54	Yes	1.5
	9	NM_016474-Ex9-F NM_016474-Ex9-R	agaggtggcctgaggatcaa tcctgccagtacctcaaagg	389	TD54	No	1.5
	10	NM_016474-Ex10-F NM_016474-Ex10-R	ctgcctgggacatgttaggt cacaaggcagcacgaagtta	387	TD54	Yes	1.5
	11	NM_016474-Ex11a-F NM_016474-Ex11a-R	agttttgctgttgccattctt aatccaacccgaagtctgtg	572	TD54	Yes	1.5
		NM_016474-Ex11b-F NM_016474-Ex11b-R	cacgctgacttgggtttgta tgtgattacttccgggtcct	494	TD54	Yes	1.5
		NM_016474-Ex11c-F NM_016474-Ex11c-R	ggaccataggtcacgaggaa ggcacactgctgactttctg	481	TD54	No	1.5
		NM_016474-Ex11d-F NM_016474-Ex11d-R	attcaggagcctgacttgga aagtctgagatctgccccta	474	TD54	Yes	1.5
		NM_016474-Ex11e-F NM_016474-Ex11e-R	agtgccccatcctcattaca tcgtagctttgtggtcaatctt	466	TD54	Yes	1.5
C3orf20	1	NM_032137-Ex1-F NM_032137-Ex1-R	tgtgacatgcagatgaggtag caggatgactgcagtagtttca	250	54	No	1.5
	2	NM_032137-Ex2-F NM_032137-Ex2-R	gatggtgcatcttttgatcact atggttccaggaaagcttca	379	54	No	1.5
	3	NM_032137-Ex3a-F NM_032137-Ex3a-R	ctgatggtgtctctgggtagg ACAAAGGTGGGTTCCATGAG	487	54	Yes	1.5
	3	NM_032137-Ex3b-F NM_032137-Ex3b-R	CATCAGTGACCCTTCAGTGC gttcccagcatgtgcctct	398	54	No	1.5
	4	NM_032137-Ex4-F NM_032137-Ex4-R	ttcctgggccatttacagag ctttccattgcattcctgct	294	54	No	1.5
	5	NM_032137-Ex5-F NM_032137-Ex5-R	caatgcggcactttacagtatg gcagagccgctatgtatcct	380	54	No	1.5
	6	NM_032137-Ex6-F	catccaggtgacactgatgc	297	54	Yes	1.5

		NM_032137-Ex6-R	tccactcacatatacacagggact				
	7	NM_032137-Ex7-F NM_032137-Ex7-R	ggcaagggctattcttagacg tgaactgaacgtgaggcttg	697	54	Yes	1.5
	8	NM_032137-Ex8-F NM_032137-Ex8-R	gctaggtcatcgttgggttc ctgcagaaacagcctcagc	497	54	No	1.5
	9	NM_032137-Ex9-F NM_032137-Ex9-R	agcatcagcaggtgttaagg tgggcttttattgctaggc	400	54	Yes	1.5
	10	NM_032137-Ex10-F NM_032137-Ex10-R	tcttctccatgaagcggaat taagagcaaaggccctgaga	378	54	Yes	1.5
	11	NM_032137-Ex11-F NM_032137-Ex11-R	ctctgaaccggagcaagtct tctgtcccttgagatgacctg	300	54	Yes	1.5
	12	NM_032137-Ex12-F NM_032137-Ex12-R	gaggtgccggagaatcaat ttgggttcactatcctgtgc	464	54	No	1.5
	13	NM_032137-Ex13-F NM_032137-Ex13-R	gcctttgaccagcaggag gcaggtgctcagccagtc	494	54	Yes	1.5
	14	NM_032137-Ex14-F NM_032137-Ex14-R	tgtacagggctccaggactc gtgctgggtgactcttggtc	400	54	Yes	1.5
	15	NM_032137-Ex15-F NM_032137-Ex15-R	ctgtggcgctgttctagtga gtggaggcttcccacctc	395	54	Yes	1.5
	16&17	NM_032137-Ex16&17a-F NM_032137-Ex16&17a-R	aggctcccagaacactgatg atgccgaatgtgtgattcat	585	54	Yes	1.5
		NM_032137-Ex16&17b-F NM_032137-Ex16&17b-R	tgcacctgttaaaggccact acgtgacctgggacaagtct	592	54	Yes	1.5
FGD5	1	NM_152536-Ex1a#2F NM_152536-Ex1a#2R	catcccaaagtgctctgagtc CGTGGCCTCTTCACATCC	646	TD54	Yes	1.5
		NM_152536-Ex1bF	GATGCTGAGGACACCAGTGA	673	TD54	Yes	1.5

	NM_152536-Ex1bR	CACCCTCTGCCTCACAGTTC				
	NM_152536-Ex1cF NM_152536-Ex1cR	AGGAAGAACACCAGCACGAG TCCGGTCAACTTCCAGAATC	670	TD54	Yes	1.5
	NM_152536-Ex1dF NM_152536-Ex1dR	GCATGTGGATGTGAACGTGT gtcccatttcacggatgaag	700	TD54	Yes	1.5
2	NM_152536-Ex2F NM_152536-Ex2R	atgaatgggacccctggtag caccaagggaaggagcagat	298	TD54	No	1.5
3&4	NM_152536-Ex3&4F NM_152536-Ex3&4R	gcttggaacttgcctccag ccaggcctctgttaatggaa	400	TD54	No	1.5
5&6	NM_152536-Ex5&6F NM_152536-Ex5&6R	ggtgcatgtggtcatctctg gatttgcattaagcccttcg	839	TD54	No	1.5
7	NM_152536-Ex7F NM_152536-Ex7R	gctgctgtgagcatggtg agaggcaaggtcaccacagt	249	TD54	Yes	1.5
8	NM_152536-Ex8F NM_152536-Ex8R	gatgcttcattcttccaactct tcctgggccttggaattg	250	TD54	Yes	1.5
9	NM_152536-Ex9F NM_152536-Ex9R	gaggcctgtgacctgacact ctaggagtggggcatcagaa	233	TD54	No	1.5
10	NM_152536-Ex10F NM_152536-Ex10R	gatcccagtcctgccttgta tccatgggtagacagcacct	346	TD54	Yes	1.5
11	NM_152536-Ex11F NM_152536-Ex11R	atgatccaagccctgctcta tcattcaaataacaaatgagtggtg	500	TD54	No	1.5
12	NM_152536-Ex12F NM_152536-Ex12R	gggatcctttgaggacagaa cttgctgtagcccaacactg	299	TD54	Yes	1.5
13	NM_152536-Ex13F NM_152536-Ex13R	gtggacatggtccctctgtt ttctgaaccttcccatgctc	241	TD54	No	1.5
14	NM_152536-Ex14F NM_152536-Ex14R	aaggggcatctgagtgagaa gtgtgagtggcaacagcaat	277	TD54	Yes	1.5
15	NM_152536-Ex15F NM_152536-Ex15R	aggagaggcctttcacatca cctcactctggggcttacag	396	TD54	Yes	1.5

	16	NM_152536-Ex16F NM_152536-Ex16R	aaagtaggggacacgcacag tgaaccatggaccacagaga	365	TD54	Yes	1.5
	17	NM_152536-Ex17F NM_152536-Ex17R	acagcccagtccctgctac ctgactccgtctacccagga	461	TD54	Yes	1.5
	18	NM_152536-Ex18F NM_152536-Ex18R	gaccttatccagtgtcaaagca aagacttgactgccaccagaa	391	TD54	Yes	1.5
	19	NM_152536-Ex19F NM_152536-Ex19R	ccacacagaggcactcttca agttctgcagttctgctatgg	396	TD54	Yes	1.5
	20	NM_152536-Ex20aF NM_152536-Ex20aR	tccaggccttttgctgaata ACAGCCAGGTCTGCACAAAT	817	TD54	Yes	1.5
		NM_152536-Ex20bF NM_152536-Ex20bR	GCTGCAACCTCCAATTTTGT tgcaagtaagtccagagatgg	839	TD54	Yes	2.5
NR2C2	1	NM_003298-Ex1#2-F NM_003298-Ex1#2-F	gggtcacgaactctgaccttt ctccgtcgagaacaaaatgg	586	TD54	Yes	1.5
	2	NM-003298-Ex2-F NM-003298-Ex2-R	gccacagtgaaagcaagagg cccaaaaggcacaactcttt	550	TD54	Yes	1.5
	3	NM-003298-Ex3-F NM-003298-Ex3-R	caaagggatattgcttataaagtgc tggctttttgattcagcttg	296	TD54	Yes	1.5
	4	NM-003298-Ex4-F NM-003298-Ex4-F	ggtttttggcagtcaccact gcctcgaaggcagatgtaag	375	TD54	Yes	1.5
	5	NM-003298-Ex5-F NM-003298-Ex5-F	ctcaagtgatccacccacct ctaccaccccacctcacaac	298	TD58	Yes	1.5
	6	NM-003298-Ex6-F NM-003298-Ex6-F	tgagacttgcacaggaactga caaaatgaagcctgagtgcat	400	TD54	Yes	1.5
	7	NM-003298-Ex7-F NM-003298-Ex7-F	ccctgcagtaatggatttgg aacacaacgctggaggctat	358	TD54	Yes	1.5
	8	NM-003298-Ex8-F NM-003298-Ex8-F	tgtggaaggacaggtaaatgc tctctgaaggccaggaagaa	275	TD54	Yes	1.5
	9	NM-003298-Ex9-F NM-003298-Ex9-F	gctgtttttgaaccagcatt gctgttttccaccctcacaa	300	TD54	Yes	1.5

10	NM-003298-Ex10-F NM-003298-Ex10-F	gtgaaaccccatctctacgg aaaactgcagatcggctttg	386	TD54	Yes	1.5
11	NM-003298-Ex11-F NM-003298-Ex11-F	agcatgacagggatacatgttg caaactagacaggccctcaga	299	TD54	Yes	1.5
12	NM-003298-Ex12-F NM-003298-Ex12-F	tgggaaatgctgggacttag gctgtgggaaactgaggcta	386	TD54	Yes	1.5
13&14	NM-003298-Ex13&14a-F NM-003298-Ex13&14a-R	aatggtcccatttctttcca cctcttcatctcctgcgaac	684	TD58	Yes	1.5
14	NM-003298-Ex13&14b-F NM-003298-Ex13&14b-R	agactacaccatacaaaagtgtgc tcaaagactgatctagtttgggta	600	TD54	Yes	1.5
	NM-003298-Ex13&14c-F NM-003298-Ex13&14c-R	gccagatgtccagaaccaat tgtgatccaatggaggctct	497	TD54	Yes	1.5
15	NM-003298-Ex15a-F NM-003298-Ex15a-R	aggagcctttgctgagctgt CACATACTATTCTCTCAGGTTCTTTCA	596	TD58	Yes	1.5
	NM-003298-Ex15b-F NM-003298-Ex15b-R	GCCTTTTGATGAAGCAGCAG GAACCCCAGGTTCTAGAGGAG	593	TD58	Yes	1.5
	NM-003298-Ex15c-F NM-003298-Ex15c-R	TGTGTGCACATGTTGTGGAG CTTGCCAACAGAGATGCTGA	685	TD54	Yes	1.5
	NM-003298-Ex15d-F NM-003298-Ex15d-R	TGCAGCTCCATCATAACTGTG ACCTGTGAGTCAGGGGAGTG	599	TD54	Yes	1.5
	NM-003298-Ex15e-F NM-003298-Ex15e-R	GAACAGTGGCATGTGGAAGA AAGGATCAAAACGACGGAA	673	TD54	Yes	1.5

	G				
NM-003298-Ex15f-F NM-003298-Ex15f-R	CCCTGATGTGTTTCAGTTTCA AAGCCCTGTGAAAGTTGGAG	697	TD54	Yes	1.5
NM-003298-Ex15g-F NM-003298-Ex15g-R	TGGTTTCTGAGATCCTCTTGG AAGCCCAGAACAGCAAAAG A	691	TD54	Yes	1.5
NM-003298-Ex15h-F NM-003298-Ex15h-R	TGAAAGACCAAATTAGTTGA GCA GATGAGACTTCCTATGACCT GGA	700	TD54	Yes	1.5
NM-003298-Ex15i-F NM-003298-Ex15i-R	AATCAGCAGTTCATGCCACA TCACCCAAAGGAAAATGGAG	700	TD54	Yes	1.5
NM-003298-Ex15j-F NM-003298-Ex15j-R	GCATTGTGCGTGTACAGGTT TTTAACTGGCGATAGAAAGA TGG	693	TD58	Yes	1.5
NM-003298-Ex15k-F NM-003298-Ex15k-R	CGTCACGGTCAGAGATTCAG CAGGGAAGCTGAAGGCTAAA	661	TD54	Yes	1.5
NM-003298-Ex15l-F NM-003298-Ex15l-R	GCTCATAGACCTGGGAAGCA AGCAGCATTTGAGTCCATGA	657	TD58	Yes	1.5
NM-003298-Ex15m-F NM-003298-Ex15m-R	TGTTCAACATTTATTGATTGA TAGACT aacacaagagagcgctgcat	828	TD58	Yes	1.5
BC051670-Ex1-F BC051670-Ex1-F	ccctaagctttgctgattgg tgagcctcattaggggattc	370			
AK094590-Ex1a-F AK094590-Ex1a-F	cctaccccattccttcctgt cagagcaagaccctgtctca	557			
AK094590-Ex1a-F AK094590-Ex1a-F	aatcatttctgtttatgacctttc ggggaaaggaggagtgacta	600			
AK094590-Ex1a-F AK094590-Ex1a-F	ccactacccctggctttctt cagagcgagactccatctca	596			

		AK094590-Ex1a-F AK094590-Ex1a-F	tgccaaataatgcacctctg aacacaaaattggccaggag	490			
MRPS25	1	1-NM_022497-F 1-NM_022497-R	gcctacaagtcccagagtgc gactcgggacctcacgttac	400	TD54	No	1.5
	2	2-NM_022497-F 2-NM_022497-R	tgcagtacaggttgggataca aaaggtgccaggagtcacag	467	TD54	No	1.5
	3	3-NM_022497-F 3-NM_022497-R	tgcgctgtgagttctgtttc caagctgagattcccaggag	457	TD54	No	1.5
	4a	4a-NM_022497-F 4a-NM_022497-R	ttgtctcggctgagaaactg tgtggccttaacctctcagg	583	TD54	No	1.5
	4b	4b-NM_022497-F 4b-NM_022497-R	atatccctgctctgctggtc aagtgacagtaacatcagtgtcctatg	688	TD54	No	1.5
	4c	4c-NM_022497-F 4c-NM_022497-R	GGTGTCTATAGCTGGCTCTG C CCTGGAAAGCTTTTAGGACT TG	700	TD54	No	1.5
	4d	4d-NM_022497-F 4d-NM_022497-R	cacggacccctagaaactca cactgcagatgaaagccaaa	663	TD54	No	1.5
	4e	4e-NM_022497-F 4e-NM_022497-R	TCTTGACCTTGATACCACCTG TGCTTGTGCAAGTAAAATTA AGAA	597	TD54	No	1.5
	4f	4f-NM_022497-F 4f-NM_022497-R	ccctgtaggctacttgatcctg ccctggcaggacagtcttta	681	TD54	No	1.5
	4g	4g-NM_022497-F 4g-NM_022497-R	ggcctgtctgtaatggcatc cagctttcttctccctgtgaa	599	TD54	No	1.5
	4h	4h-NM_022497-F 4h-NM_022497-R	TTCAAGACTGAAGACATTGA TTAGA aaagtatctttgcggggagaa	693	TD54	No	1.5
ZFYVE20	1	NM_022340-Ex1-F NM_022340-Ex1-R	ctcgacaggcgtctaccag cgtctccgttctggaggagt	236	TD54	No	1.5

2	NM_022340-Ex2-F NM_022340-Ex2-R	gtccaattcggtggttgtgt atgcactcacccctaacgac	397	TD54	No	1.5
3	NM_022340-Ex3-F NM_022340-Ex3-R	cgcccattagagataagtgctc tagaccaagcaagctgcaaa	437	TD54	No	1.5
4	NM_022340-Ex4-F NM_022340-Ex4-R	gcttgcttggtctaaggatttg gtggcaacaccaaagctgat	561	TD54	No	1.5
5	NM_022340-Ex5-F NM_022340-Ex5-R	tgcttgctcctgcattttac gcctcctcagccaactctta	384	TD54	No	1.5
6	NM_022340-Ex6-F NM_022340-Ex6-R	tgttgaccaacctgcttttg cccaggctagaactgtccac	377	TD54	No	1.5
7&8	NM_022340-Ex7&8-F NM_022340-Ex7&8-R	tttgtgattttgtggcgttg tgaatgaagattcccatgagc	576	TD54	No	1.5
9	NM_022340-Ex9-F NM_022340-Ex9-R	cacagctggagcctgatctt cgaacaacttcccaaagcat	469	TD54	No	1.5
10	NM_022340-Ex10-F NM_022340-Ex10-R	acaattaatgtgttagaaaagtgaaca gatggtagcctgtaccccact	395	TD54	No	1.5
11	NM_022340-Ex11-F NM_022340-Ex11-R	gtctgaaatgccacctttagaga gagtgcaggtctgctttcg	298	TD54	No	1.5
12	NM_022340-Ex12-F NM_022340-Ex12-R	catgaaagcctggcaactc aaacaagtcattggcatgga	287	TD56	Yes	2.5
13	NM_022340-Ex13-F NM_022340-Ex13-R	gagaacagagtggctcagctt actacaaccgcacacaacct	300	TD54	No	1.5
14	NM_022340-Ex14a-F NM_022340-Ex14a-R	gcacagggatagagggacaa TCCCTTTCTCGTTCCAACTC	492	TD56	Yes	2.5
	NM_022340-Ex14b-F NM_022340-Ex14b-R	GACGAGTATGACCAGCAGCA GTCCATCTCAAAGGGGTTGA	690	TD54	No	1.5
	NM_022340-Ex14c-F NM_022340-Ex14c-R	CAGACAGCCCAGCTCCTAAC TAGCAAATGGACCCAAGAGG	687	TD54	No	1.5
	NM_022340-Ex14d-F NM_022340-Ex14d-R	TCAGCATTCAGTTGCTGCTT CGGCCACCACTCTTATTTTC	566	TD54	No	1.5

		NM_022340-Ex14e-F NM_022340-Ex14e-R	AGGTGGGGTGATGGGAATA GGCTGATCAAATAGCCTTGG	697	TD54	No	1.5
		NM_022340-Ex14f-F NM_022340-Ex14f-R	GCCGGGTCTCTAATTTAGGA AT AAATTAACAATTTCCCTTTCT GG		TD54	No	1.5
		NM_022340-Ex14g-F NM_022340-Ex14g-R	CCGCCCTGATGAGATATTTT AGAAATGTCATGGGCCTCAG	699	TD54	No	1.5
		NM_022340-Ex14h-F NM_022340-Ex14h-R	AACCCTGCTGGTTGTTATGG AGCGATTCTCGTGCCTCAG	599	TD58	Yes	1.5
		NM_022340-Ex14i-F NM_022340-Ex14i-R	TCCTGTAATCCCAGCACTTTG CATCTGATGCCTGAATCCCT A	697	TD58	Yes	1.5
		NM_022340-Ex14j-F NM_022340-Ex14j-R	GGTTGACAGCAGGCAAACTT ttgtaacacatgtgcacacagac	581	TD54	No	1.5
CAPN7	1	NM_014296-Ex1-F NM_014296-Ex1-R	gccctcagctgctcaatc gacgagggaaaagcaaagc	678	TD54	Yes	1.5
	2	NM_014296-Ex2-F NM_014296-Ex2-R	ttttgtcattcatggaaacca gtggaatgagcagcaacaga	339	TD56	Yes	1.5
	3	NM_014296-Ex3-F NM_014296-Ex3-R	aaatgaagaatacagggttttgg tgttaaaattcttttattcaaggtg	367	TD56	Yes	1.5
	4	NM_014296-Ex4-F NM_014296-Ex4-R	tgtttgttctcaatttcagcaag gaggtttgttactgcagccttt	286	TD54	Yes	1.5
	5	NM_014296-Ex5-F NM_014296-Ex5-R	tgcttatctttggaagtagctga tgaatacaattcagcaaacctaatg	463	TD56	Yes	1.5
	6	NM_014296-Ex6-F NM_014296-Ex6-R	tggttttgtgtcttggtgtca ccacccctctattccaacaa	498	TD54	Yes	1.5
	7	NM_014296-Ex7-F NM_014296-Ex7-R	agccagatttgaccagaagg tttccaattacacttctgtacctga	396	TD56	Yes	2
	8	NM_014296-Ex8-F	tgttgaaattacgtgaatgaatatga	242	TD56	Yes	1.5

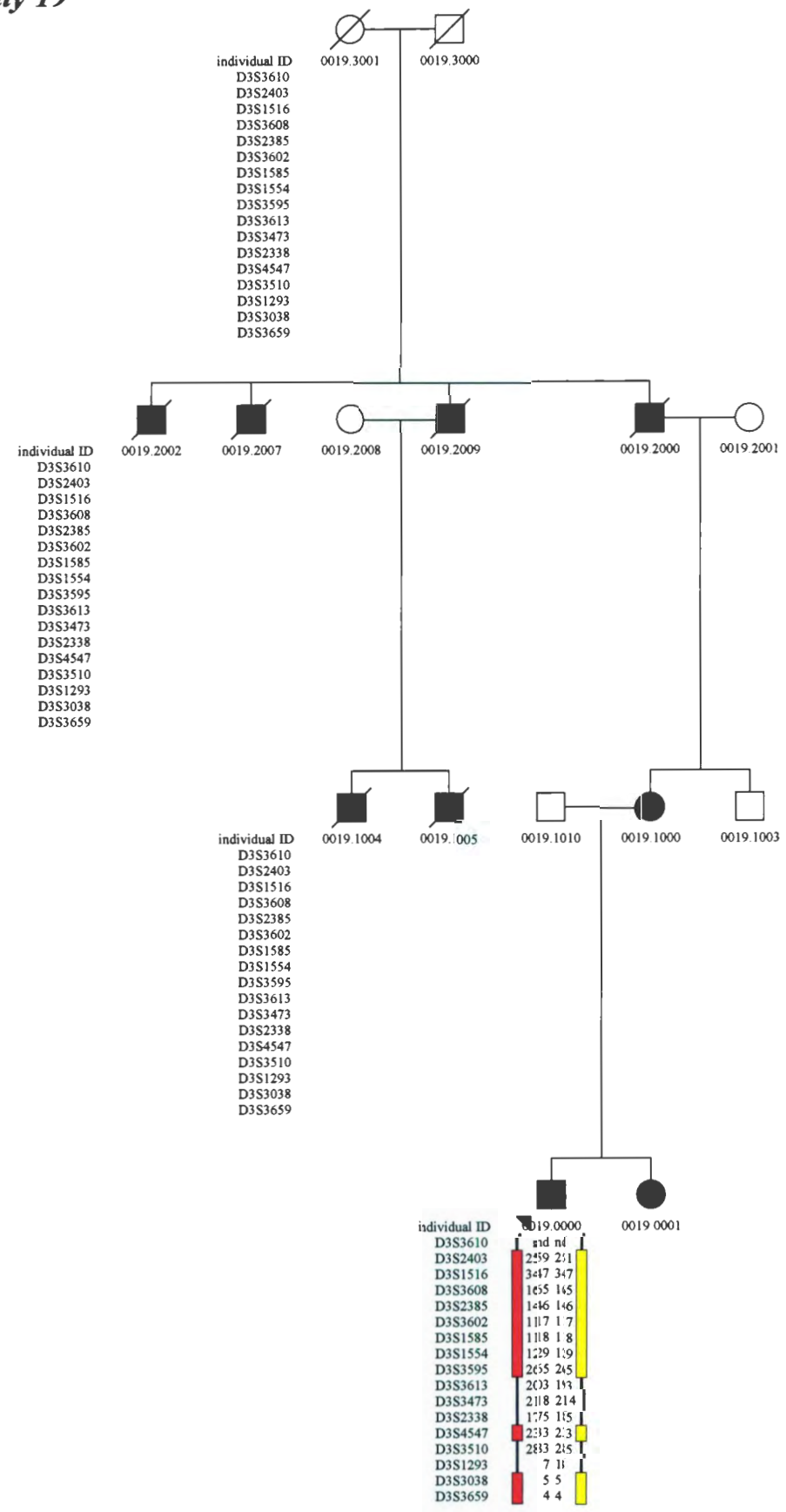
	NM_014296-Ex8-R	tgaatgatatgccctctcca				
9	NM_014296-Ex9-F NM_014296-Ex9-R	atatacctgaatgaaatgtgctgtat tgatttcactcaagattcaaagg	250	TD56	Yes	1.5
10	NM_014296-Ex10-F NM_014296-Ex10-R	cgtcagcatattgtgcgtta tcatcctgtttcaggggttg	396	TD54	Yes	1.5
11	NM_014296-Ex11-F NM_014296-Ex11-R	gctgctctggggatttaaga cttggcctcccatagtgct	283	TD54	Yes	1.5
12	NM_014296-Ex12-F NM_014296-Ex12-R	tcttcaggggtaggaacagc ttttcaacggtagagatcaagtt	397	TD56	Yes	2
13&14	NM_014296-Ex13&14-F NM_014296-Ex13&14-R	ggcagccatttacagtttttg gtcaactgatccacccacct	699	TD56	Yes	1.5
15	NM_014296-Ex15-F NM_014296-Ex15-R	ggatagcttaattcacttttgaaat cgtatgtctttcacaatgttctaact	300	TD56	Yes	1.5
16	NM_014296-Ex16-F NM_014296-Ex16-R	tttttgtggagagatttagagca aaatgtgtcatttttcaacttaatca	262	TD54	No	2
17	NM_014296-Ex17-F NM_014296-Ex17-R	cagttaaggagatgtccatcca tccagataataccacttacaaagca	332	TD54	Yes	1.5
18	NM_014296-Ex18a-AK093742-Ex1-F NM_014296-Ex18a-AK093742-Ex1-R	caggctggtctcaaactcct gcagcctagtatcaggtcagg	647	TD56	Yes	1.5
	NM_014296-Ex18#2-F NM_014296-Ex18#2-R	tgacttgtgtatgccaactgatt attcctgtgaagcctttgct	280	TD54	Yes	1.5
19	NM_014296-Ex19-F NM_014296-Ex19-R	agaggggttggagacacaga gattgtcctttcctcactcca	400	TD54	Yes	1.5
20&21	NM_014296-Ex20&21a-F NM_014296-	aaacttagtatgacatctgcacttca acaaaatggcccctaaaatg	681	TD54	Yes	1.5

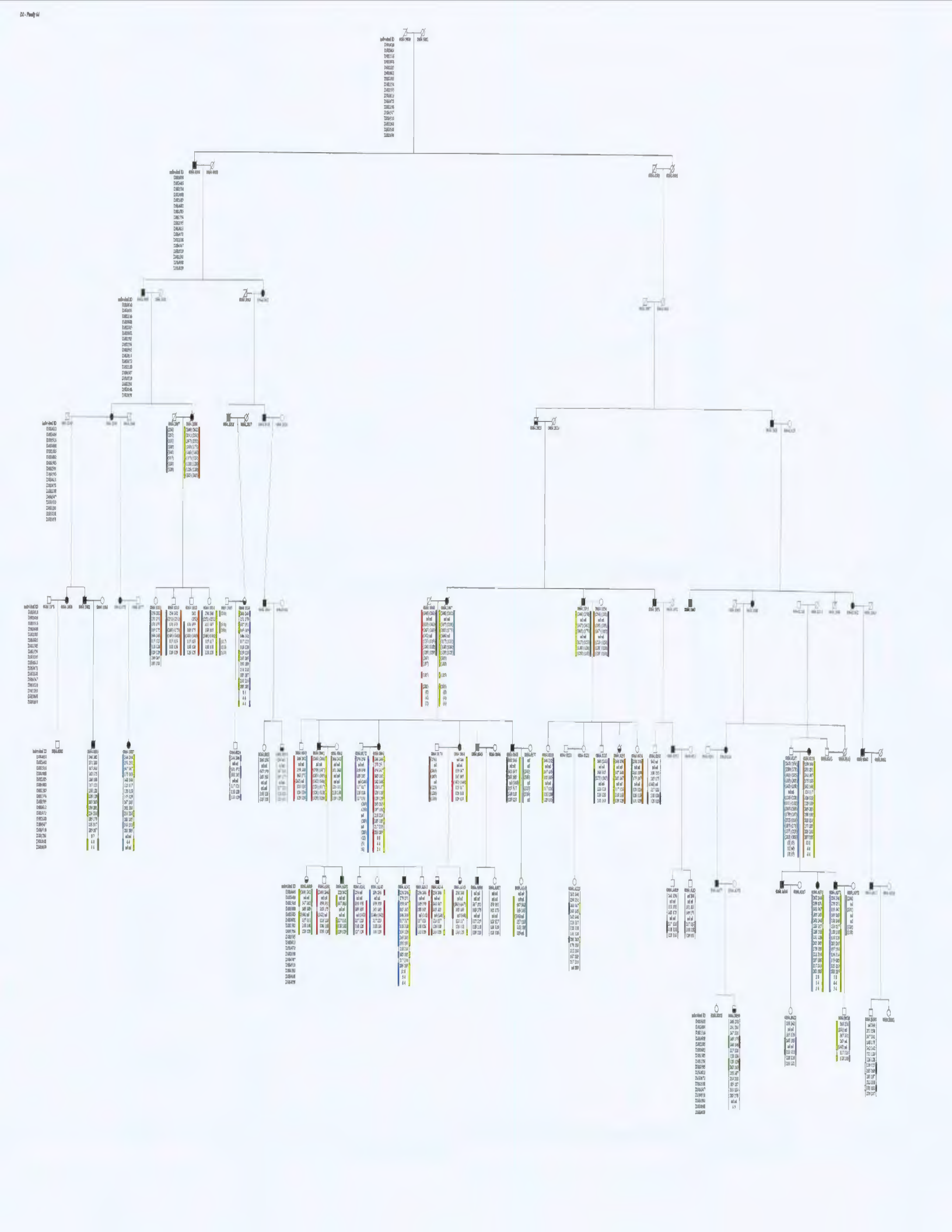
		Ex20&21a-R					
		NM_014296-Ex21b-F NM_014296-Ex21b-R	aaggacgcaaatcttcagga tgaatgggagataacctcaaca	594	TD54	Yes	1.5
		NM_014296-Ex21c-F NM_014296-Ex21c-R	caaatgaattgtgccaccataa tctaacttaaactatgcaatatccgta	581	TD54	Yes	1.5
		NM_014296-Ex21d-F NM_014296-Ex21d-R	cataaaaacccatcaggacag tttgatggcagggattaggt	491	TD56	Yes	1.5
		NM_014296-Ex21e-F NM_014296-Ex21e-R	ggtggacaaatttaacaagttttct cttgtgaaccctgggcaaa	487	TD56	Yes	1.5
	mRNA1	BI464220-Ex2-F BI464220-Ex2-R	cctcctgagtagctgggact tggtgtttccattctcctca	299	TD56	Yes	2
SH3BP5	1	NM_001018009-Ex1-F NM_001018009-Ex1-R	gacctggatagctgggactg gtacgcgtagacaccgacct	571	TD54	No	1.5
		NM_004844-Ex1-F NM_004844-Ex1-R	gactgacagttctgcattgct GCACTTTTATAAACCTAAAC TCTTCC	496	TD54	Yes	1.5
	2	NM_001018009-Ex2-F NM_001018009-Ex2-R	gtcaaaagggacaccctcaa gggtccagcaataacactcc	380	TD54	No	1.5
	3	NM_001018009-Ex3-F NM_001018009-Ex3-R	cggtgctgctctagacactg tcgagcttctaaccttcagca	385	TD54	Yes	1.5
	4	NM_001018009-Ex4-F NM_001018009-Ex4-R	aggcagagatgagcttgtcc agggtgttttcaggcaacat	390	TD54	No	1.5
	5	NM_001018009-Ex5-F NM_001018009-Ex5-R	tattccaagcctgggctgtt ggtggtccttgagtgtgtga	280	TD54	Yes	1.5
	6	NM_001018009-Ex6-F NM_001018009-Ex6-R	agtggtagctgatgcctcttg ggcccatgtgatactctgct	223	TD54	No	1.5
	7	NM_001018009-Ex7-F NM_001018009-Ex7-R	tgatgggcagaaaatgtacttg tgcccaactctgagaccttt	421	TD54	No	1.5
	8	NM_001018009-Ex8-F NM_001018009-Ex8-R	caggcttgggaagtcacatt ctggcagaggtatggaatgg	498	TD54	No	1.5

9	NM_001018009-Ex9a-F NM_001018009-Ex9a-R	gggctgagttagctgactggt TGTTCCACAAAGGCTGTGAA	596	TD54	No	1.5
	NM_001018009-Ex9b-F NM_001018009-Ex9b-R	AATGAAGTTTCAGTGACCTT GAG CAGTGGCAACCATTCTTCAC	982	TD54	No	1.5
	NM_001018009-Ex9c-F NM_001018009-Ex9c-R	GGGGAGAAGGTTCTACAGCA aagtagaattcaaaccttggtaaaca	700	TD54	Yes	1.5

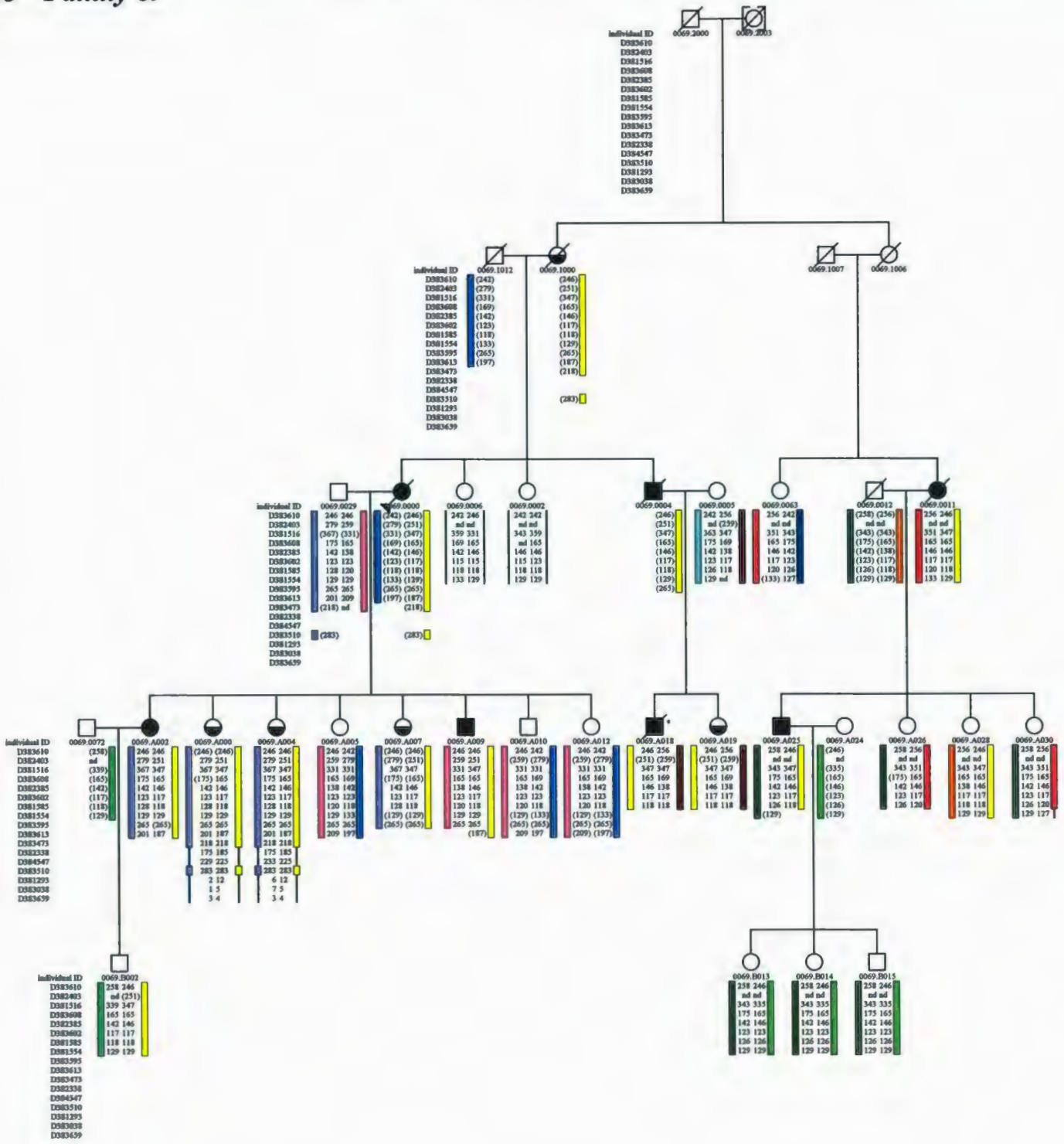
Appendix D: Haplotype construction of the ARVD5 candidate region in 15 ARVC NL families. Markers are arranged telomeric to centromeric on chromosome 3p and are shown to the left of each generation. Disease associated haplotype is colored yellow and unlinked haplotypes are various other colors. Brackets surrounding a genotype indicates that it is inferred. Least number of recombination method was used during haplotyping. Uncertain genotypes were arranged to mimic the disease-associated haplotype. Overall, haplotypes and uncertain genotypes shared a region flanked by markers D3S3610 and D3S3613.

D1 - Family 19

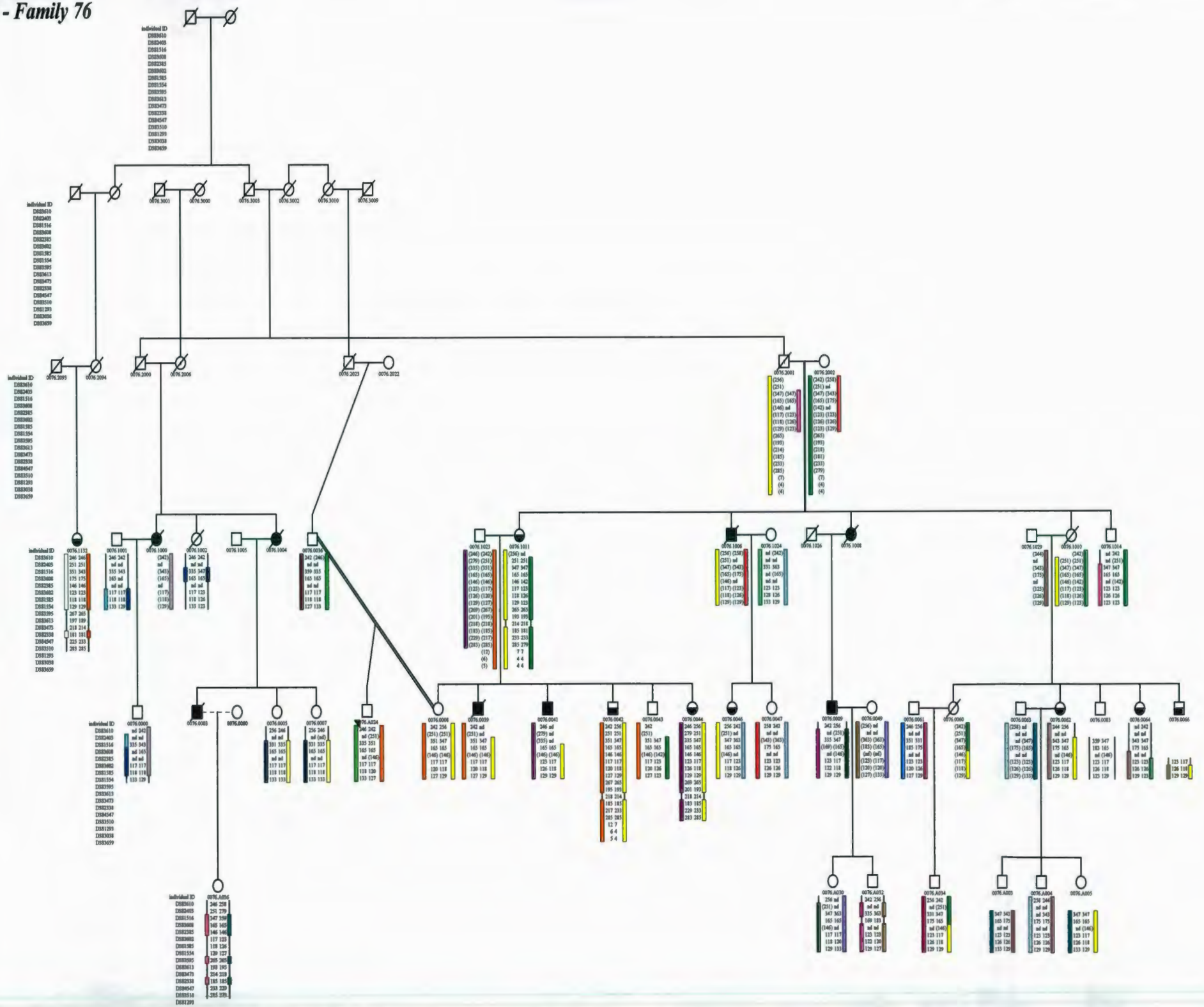

individual ID
D3S3610
D3S2403
D3S1516
D3S3608
D3S2385
D3S3602
D3S1585
D3S1554
D3S3595
D3S3613
D3S3473
D3S2338
D3S4547
D3S3510
D3S1293
D3S3038
D3S3659
0019.3001
0019.3000
0019.2002
0019.2007
0019.2008
0019.2009
0019.2000
0019.2001
0019.1004
0019.1005
0019.1010
0019.1000
0019.1003
0019.0000
0019.0001
D3S3610 nd nd
D3S2403 259 251
D3S1516 347 347
D3S3608 165 165
D3S2385 146 146
D3S3602 117 117
D3S1585 118 118
D3S1554 129 129
D3S3595 265 245
D3S3613 203 193
D3S3473 218 214
D3S2338 175 185
D3S4547 233 233
D3S3510 283 285
D3S1293 7 11
D3S3038 5 5
D3S3659 4 4



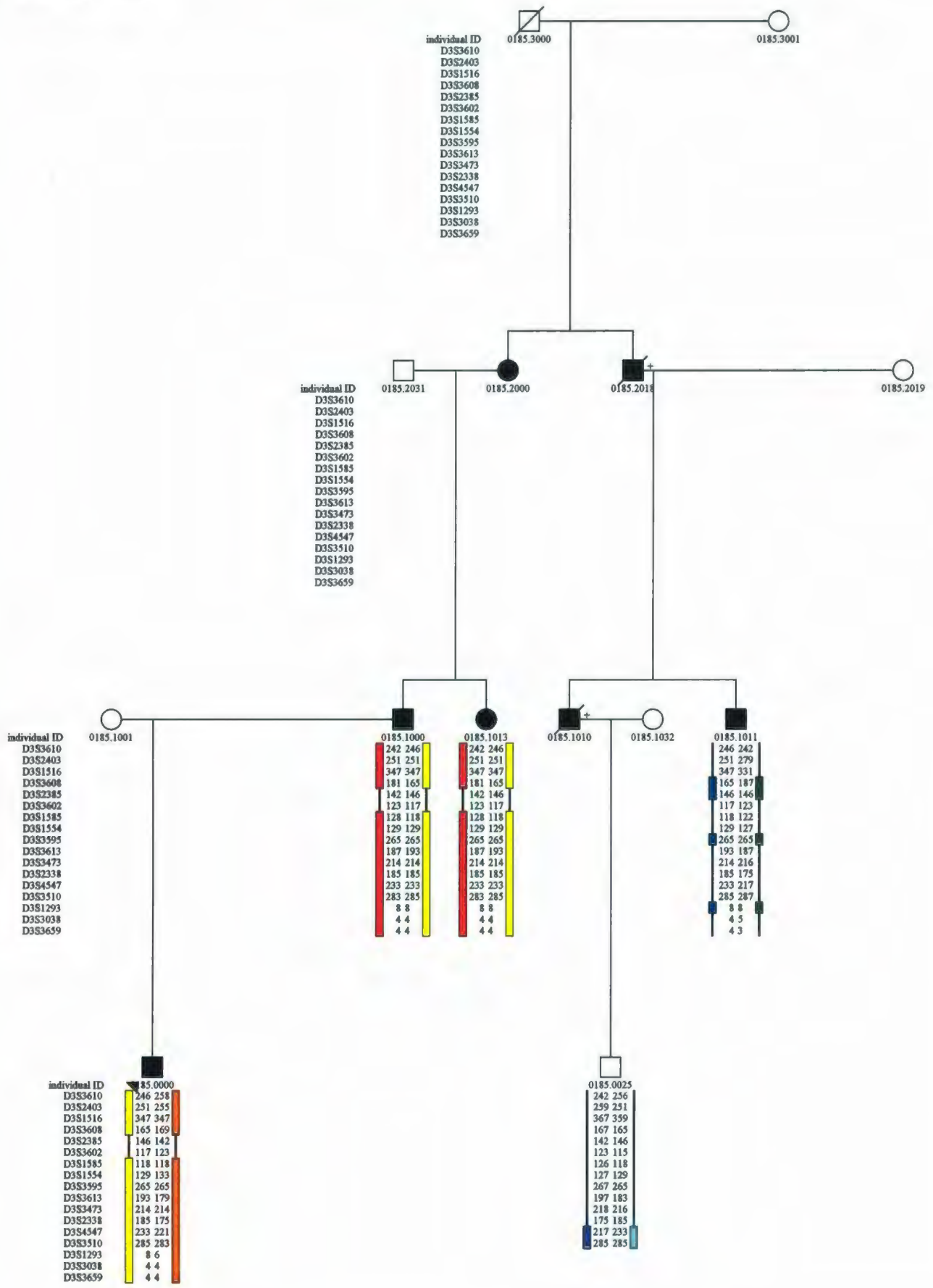
D3 - Family 69


Individual ID
D383610
D382403
D381516
D383608
D382385
D383602
D381585
D381554
D383595
D383613
D383473
D382338
D384547
D383510
D381293
D383038
D383639
0069.2000
0069.2003
0069.1012
0069.1000
0069.1007
0069.1006
0069.0029
0069.0000
0069.0006
0069.0002
0069.0004
0069.0005
0069.0063
0069.0012
0069.0011
0069.0072
0069.A002
0069.A000
0069.A004
0069.A005
0069.A007
0069.A009
0069.A010
0069.A012
0069.A018
0069.A019
0069.A025
0069.A024
0069.A026
0069.A028
0069.A030
0069.B002
0069.B013
0069.B014
0069.B015

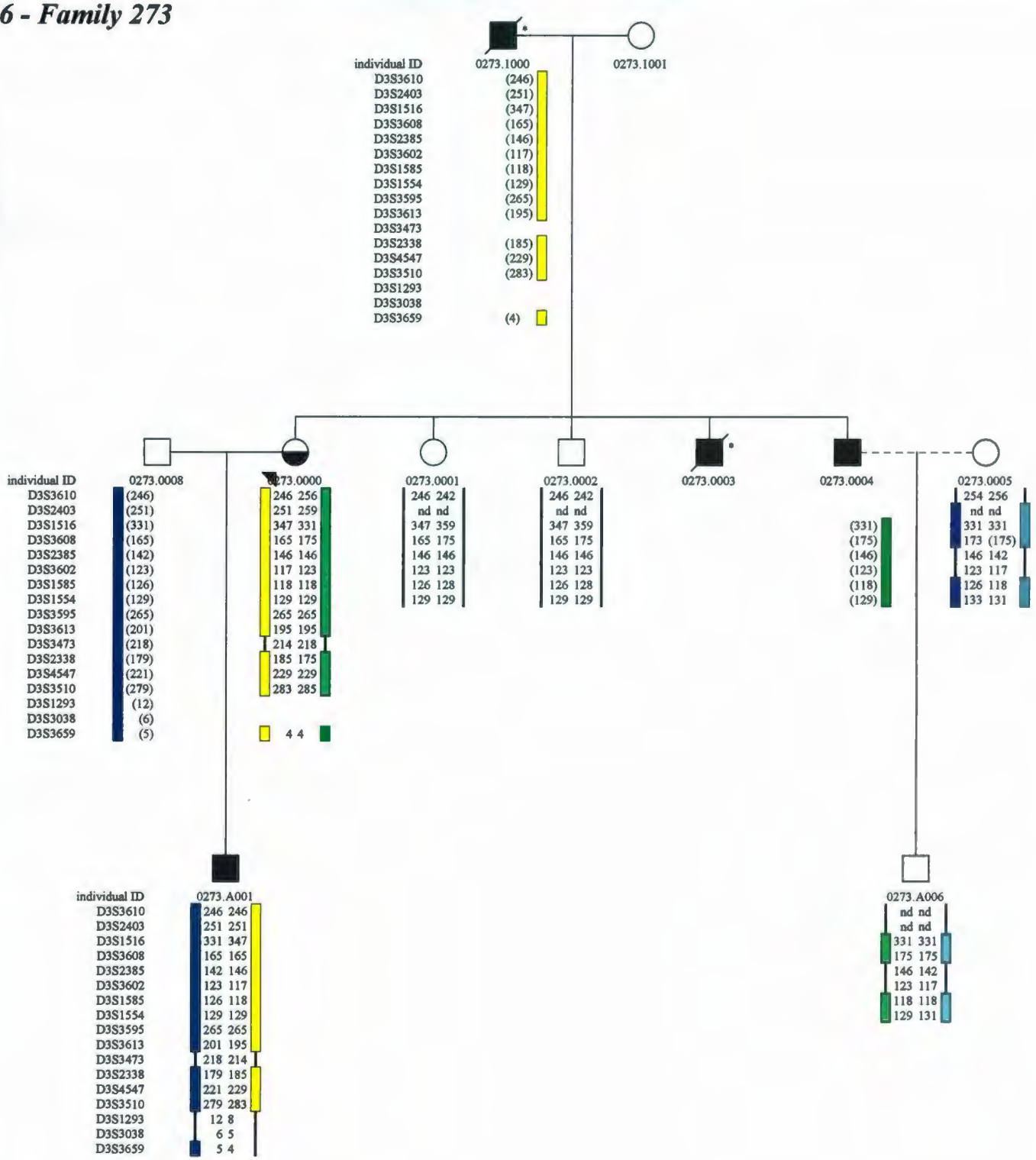
D4 - Family 76


individual ID
D883610
D882405
D881516
D883608
D882385
D883602
D881585
D881554
D883595
D883613
D883473
D882338
D884547
D883510
D881293
D883058
D883659
0076.3001
0076.3000
0076.3003
0076.3002
0076.3010
0076.3009
0076.2093
0076.2094
0076.2000
0076.2006
0076.2023
0076.2022
0076.2001
0076.2002
0076.1132
0076.1001
0076.1000
0076.1002
0076.1005
0076.1004
0076.0030
0076.1023
0076.1011
0076.1006
0076.1024
0076.1026
0076.1008
0076.1029
0076.1010
0076.1014
0076.0000
0076.0083
0076.0080
0076.0005
0076.0007
0076.A024
0076.0000
0076.0039
0076.0041
0076.0042
0076.0043
0076.0044
0076.0046
0076.0047
0076.0009
0076.0049
0076.0061
0076.0060
0076.0063
0076.0062
0076.0083
0076.0064
0076.0066
0076.A039
0076.A032
0076.A034
0076.A003
0076.A004
0076.A005
0076.A036

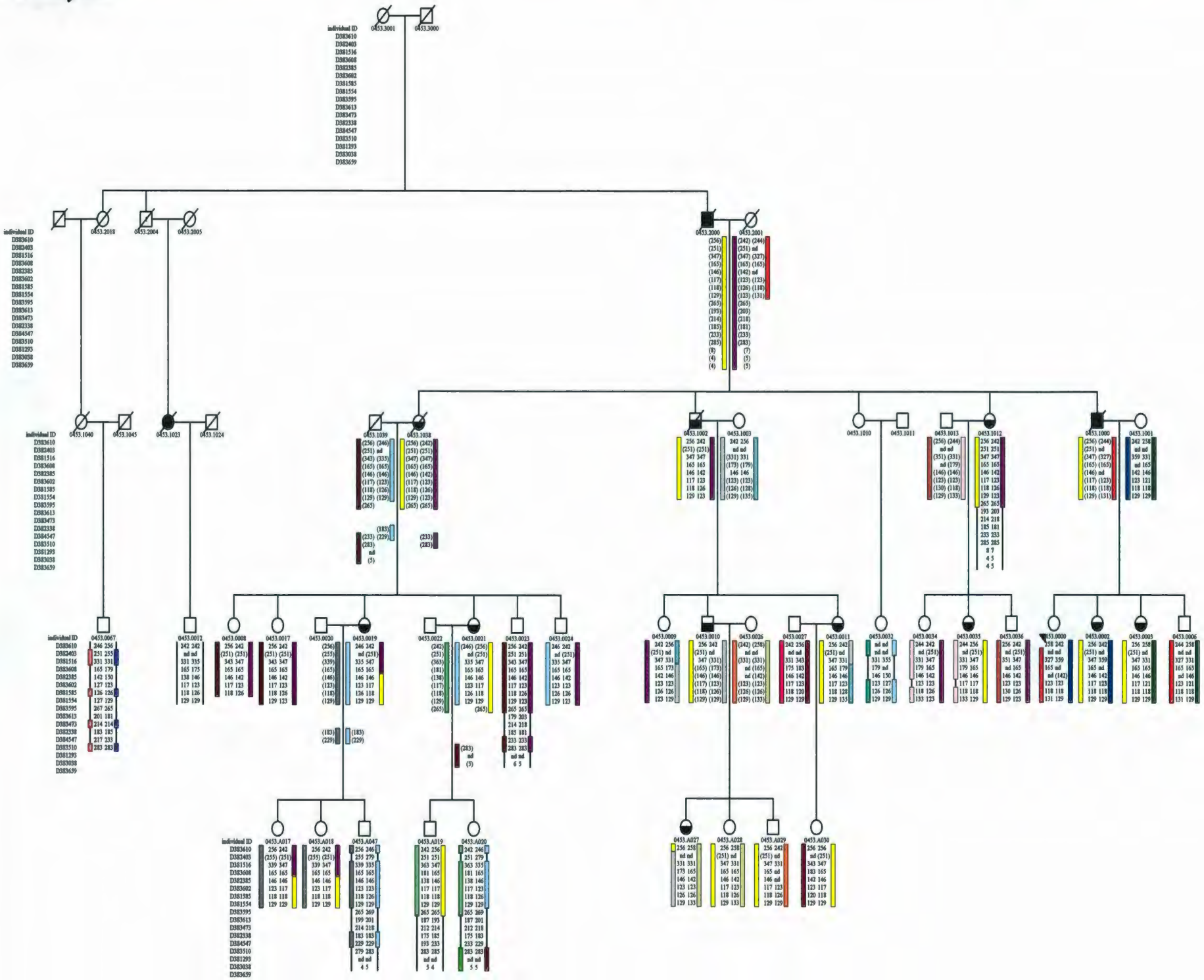
D5 - Family 185



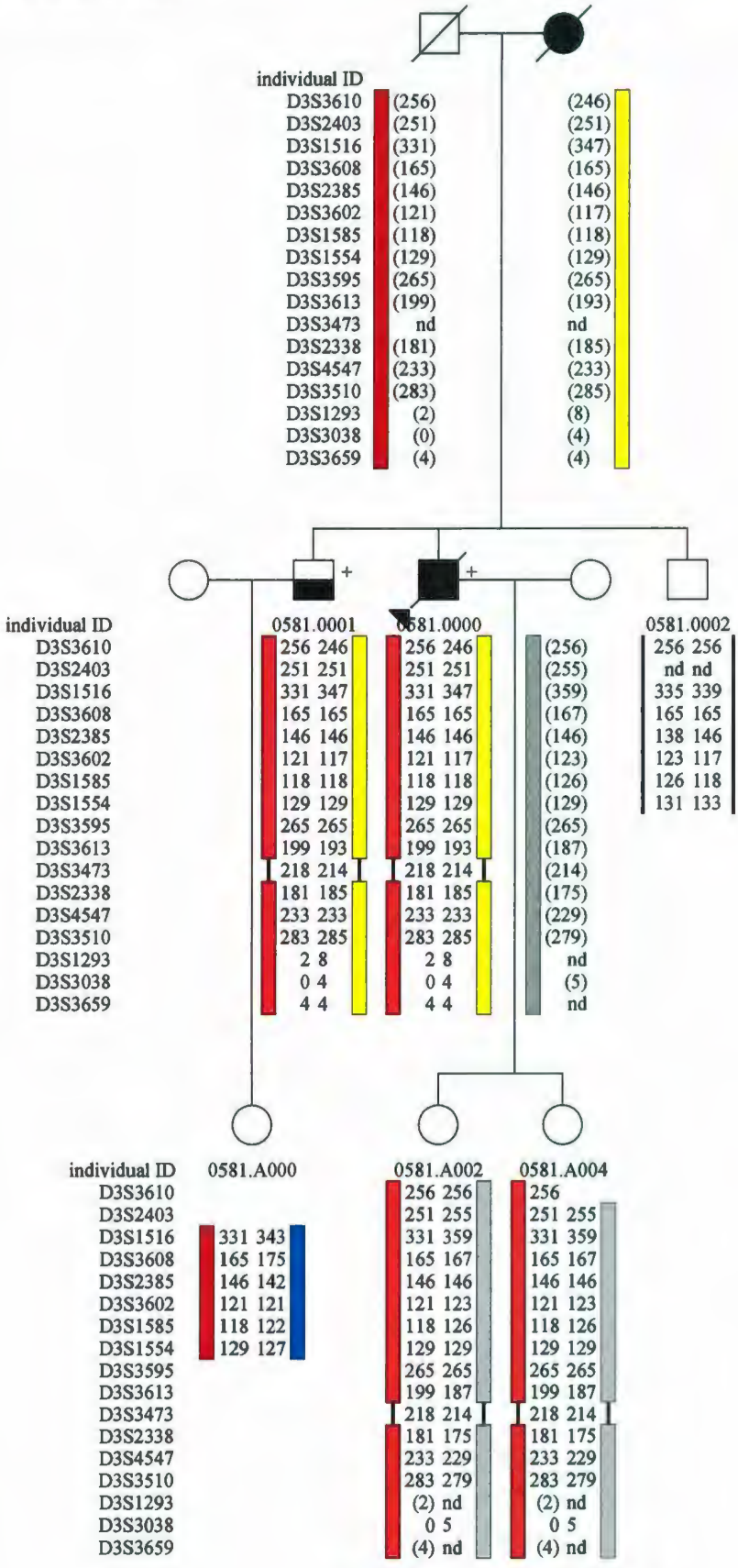
D6 - Family 273



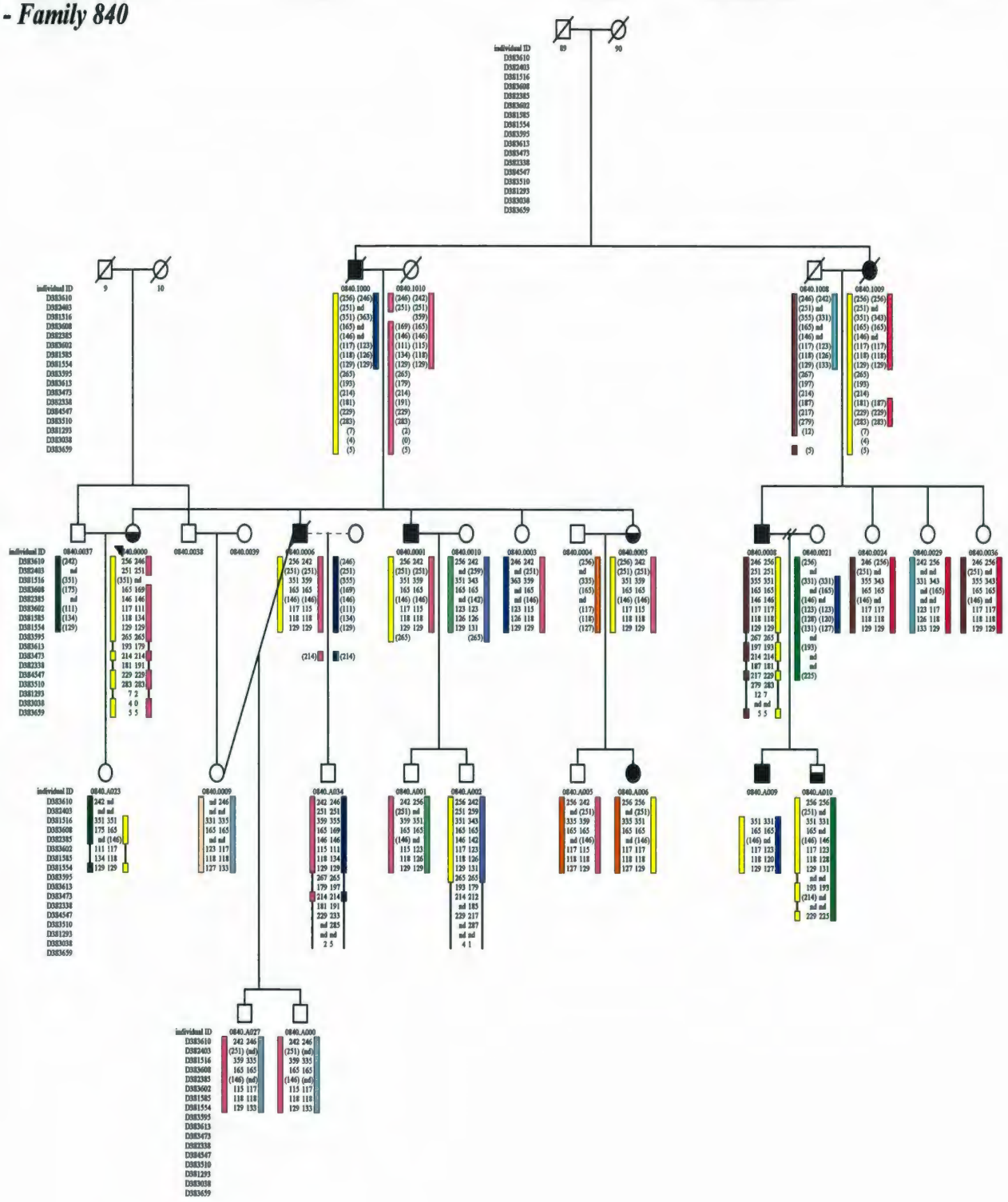
D7 - Family 453



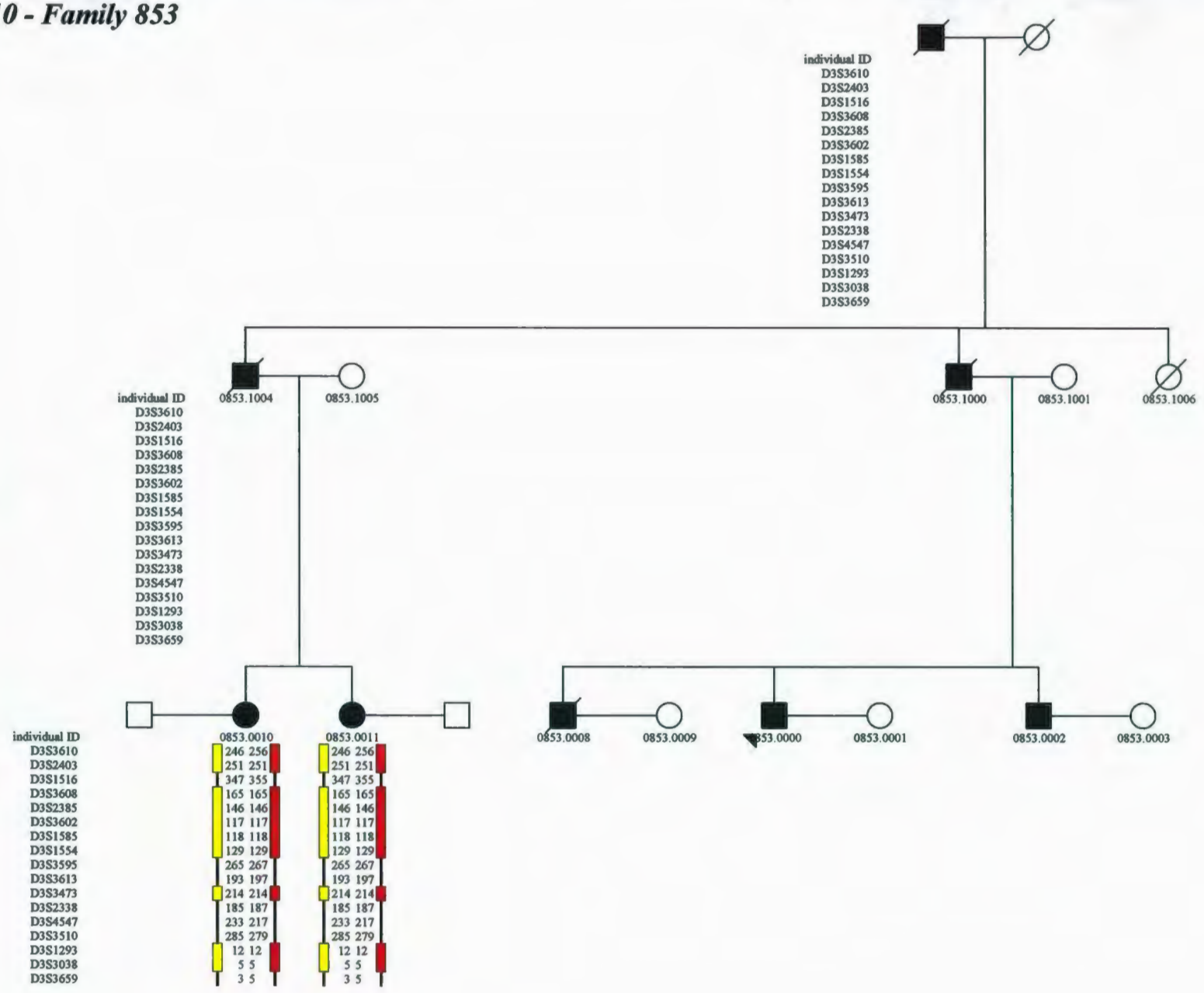
D8 - Family 581



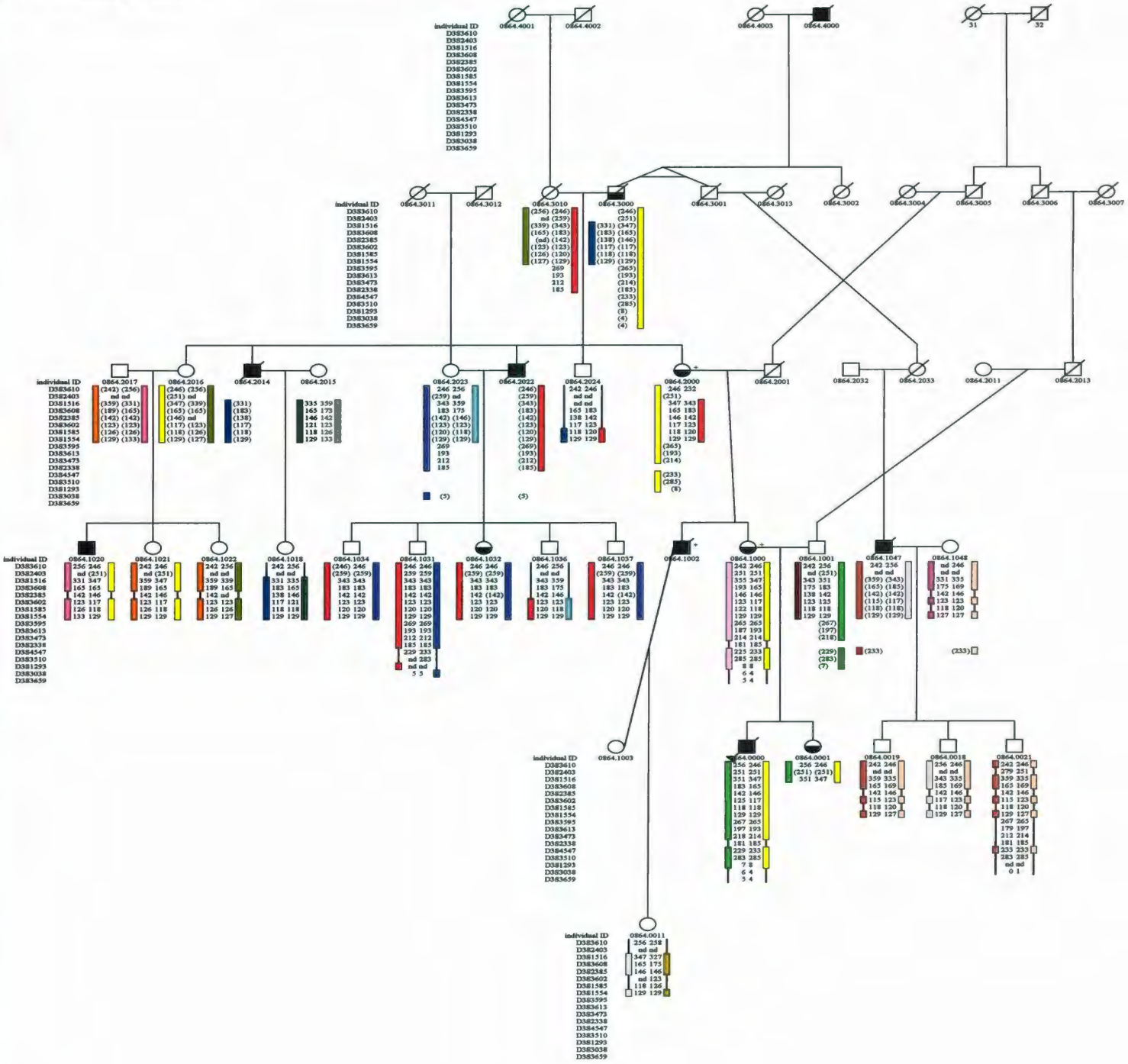
D9 - Family 840



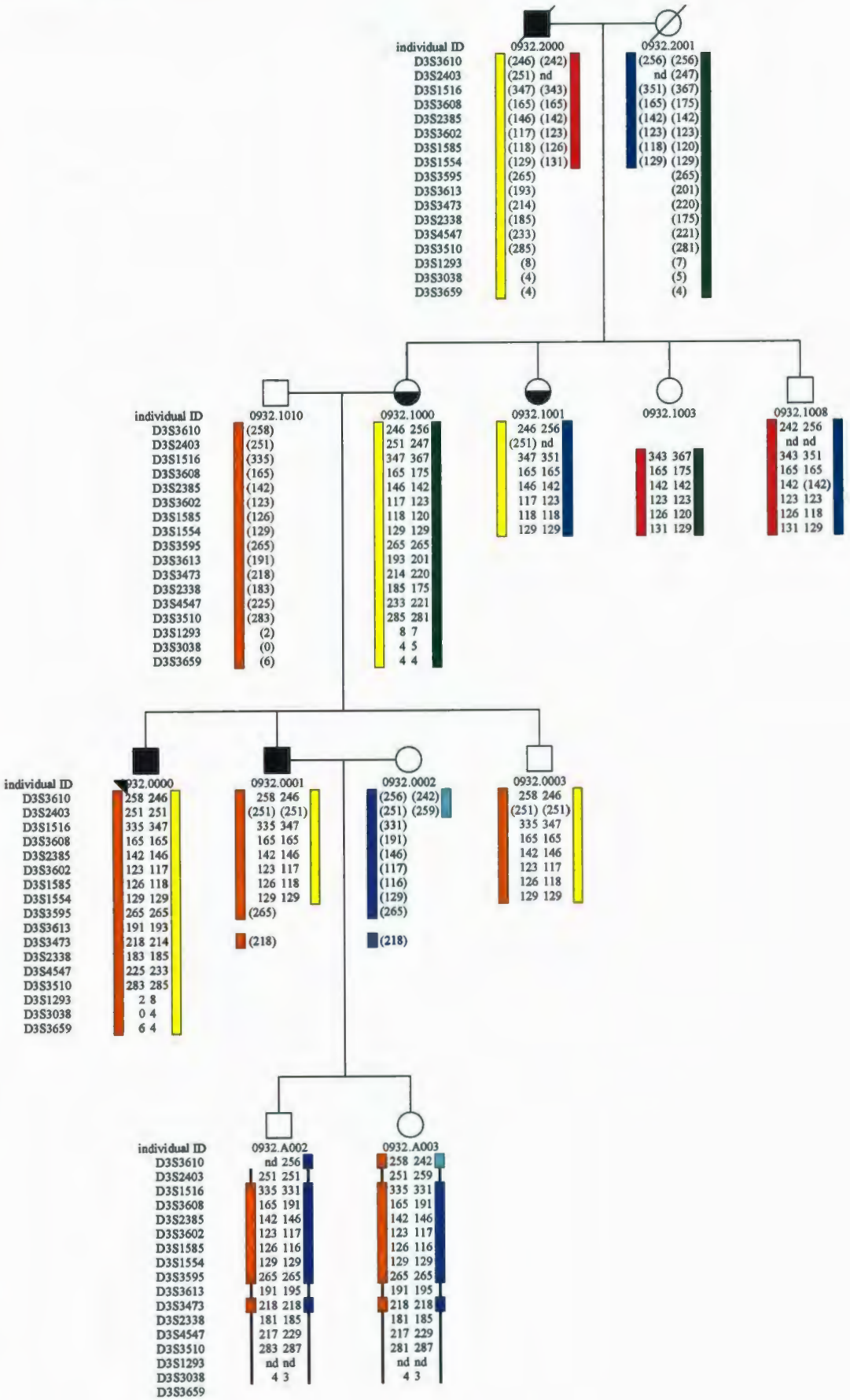
D10 - Family 853



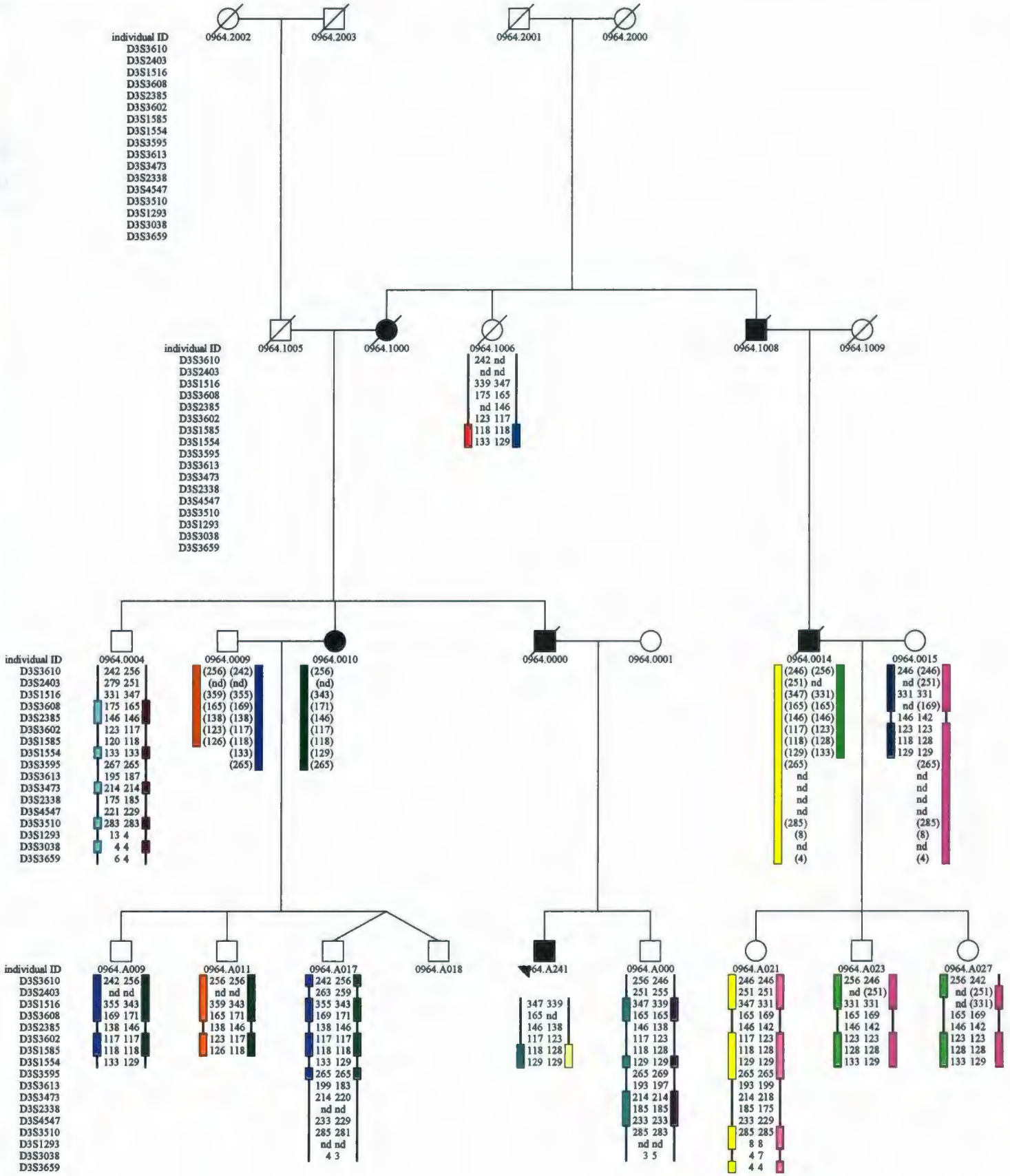
D11 - Family 864



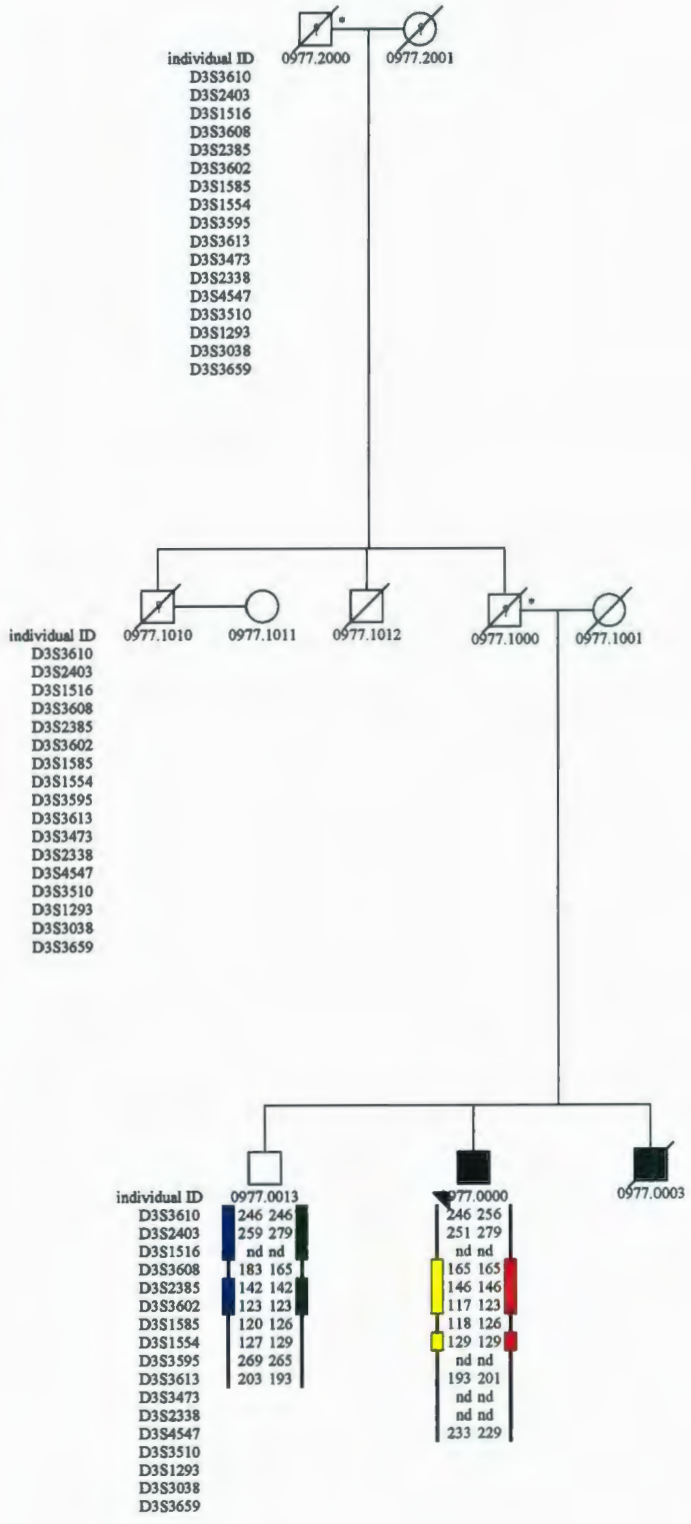
D12 - Family 932



D13 - Family 964



D14 - Family 977



D15 - Family 1139