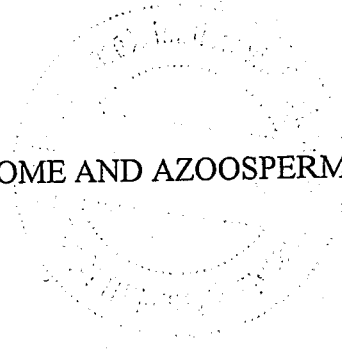


GENETIC ANALYSES IN SECKEL SYNDROME AND AZOOSPERMIA



by

Serdar Korur

B.S. in Molecular Biology and Genetics, Boğaziçi University, 2002

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Bogazici University Library



39001102534560

14

Graduate Program in Molecular Biology and Genetics
Boğaziçi University

2004

ACKNOWLEDGMENTS

I would like to express my gratitude to my thesis supervisor, Prof. Aslı Tolun for her continuous guidance and helpful advices. This study could never have been achieved without her supreme experience that she never refrained from sharing with me. Her valuable criticisms helped me in every step of my study.

I owe grateful thanks to the other members of the thesis evaluation committee, Assoc. Prof. Nur Buyru and Prof. Nihan Ünaltuna, who gave their valuable time to read, examine and evaluate this thesis.

Sibel Uğur shared her knowledge and experience all throughout my studies. She was also a great friend who was always with me during the hard times of my laboratory work. She was always giving a hand at all the steps of this study.

ABSTRACT

GENETIC ANALYSES IN SECKEL SYNDROME AND AZOOSPERMIA

The autosomal recessive genetic disorders Seckel syndrome and azoospermia were studied. In Seckel syndrome, two of the known loci were tested for linkage in 15 Seckel families. In addition, four ATR interacting and fifteen DNA repair gene loci were tested for linkage. In azoospermia, the locus which was previously identified in our laboratory was refined with additional markers, and computer based parametric tests were applied to evaluate the linkage information obtained.

The strongest candidate genes were selected using databases. *MNAT1* and *RAD51L1* were assessed as the strongest candidates for Seckel syndrome and *P2RX1* and *NYD-SP20* for azoospermia. *MNAT1*, *RAD51L1* and *P2RX1* were analyzed for mutations in relevant families by direct sequencing and SSCP, *NYD-SP20* was analyzed by direct sequencing in order to identify any disease causing defects. No mutations were found in any of the genes, but a novel polymorphism was identified in *MNAT1*.

ÖZET

SECKEL SENDROMU VE AZOSPERMİDE GENETİK ANALİZ

Otozomal çekinik bozukluklarda Seckel sendromu ve azospermide genetik incelemeler gerçekleştirildi. Seckel sendromunda bilinen üç gen bölgesine bağlantı bulunup bulunmadığı 15 ailede araştırıldı. Ayrıca ATR ile etkileşen dört gen bölgesiyle ve 15 başka DNA onarım gen bölgelerine bağlantı araştırıldı. Azospermide ise laboratuvarımızda daha önce bulunmuş gen bölgesi yeni markırlarla ayrıntılı olarak incelendi. Bulgular bilgisayar tabanlı parametrik testlerle değerlendirildi.

Daha sonra en güçlü aday genler veritabanları kullanılarak değerlendirildi. *MNAT1* ve *RAD51L1* Seckel sendromu için, *P2RX1* ve *NYD-SP20* ise azospermi için en güçlü aday genler olarak belirlendi. *MNAT1*, *RAD51L1* ve *P2RX1* dizi analizi ve SSCP kullanılarak, *NYD-SP20* dizi analizi kullanılarak hastalığa sebep olan kusuru belirlemek amacıyla ilgili ailelerde mutasyon için incelendi. Genlerde herhangi bir mutasyon bulunamadı, ama *MNAT1* geninde yeni bir polimorfizm belirlendi.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT	iv
ÖZET	v
LIST OF FIGURES	ix
LIST OF TABLES	xi
LIST OF ABBREVIATIONS	xiii
1. INTRODUCTION	1
1.1. Seckel Syndrome	1
1.1.1. Seckel Loci	2
1.1.1.1. SCKL1	2
1.1.1.2. SCKL2	3
1.1.1.3. SCKL3	3
1.1.2. Candidate Genes in SCKL2	4
1.1.2.1. <i>RBBP8</i>	4
1.1.2.2. <i>RAB31</i>	4
1.1.3. Candidate Genes in SCKL5	4
1.1.3.1. <i>MNAT1</i>	4
1.1.3.2. <i>RAD51L1</i>	5
1.1.3.3. <i>RBBP1</i>	5
1.1.4. ATR Related Proteins	5
1.1.4.1. <i>CHK1</i>	6
1.1.4.2. <i>CHK2</i>	6
1.1.4.3. <i>RAD17</i>	6
1.1.4.4. <i>RAD9</i>	7
1.1.5. Other Candidate Repair Genes	7
1.1.5.1. <i>LIG4</i>	7
1.1.5.2. <i>DNAPK</i>	7
1.1.5.3. Other DNA Repair Proteins	7
1.1.6. DNA Repair Syndromes	8
1.2. Azoospermia	9

1.2.1. Candidate Genes for Azoospermia	10
1.2.1.1. <i>P2RX1</i>	10
1.2.1.2. <i>NYD-SP20</i>	10
1.3. Families.....	10
1.4. Strategy of the Study.....	14
2. PURPOSE.....	16
3. MATERIALS	17
3.1. Blood Samples	17
3.2. Chemicals.....	17
3.3. Buffers and Solutions.....	17
3.3.1. DNA Extraction from Blood Samples.....	17
3.3.2. Polymerase Chain Reaction (PCR)	18
3.3.3. Agarose Gel Electrophoresis	18
3.3.4. Denaturing Polyacrylamide Gel Electrophoresis (PAGE)	19
3.3.5. Single Strand Conformational Polymorphism (SSCP) Gel Electrophoresis	19
3.3.6. Silver Staining	20
3.4. Oligonucleotide Primers	20
3.5. Fine Chemicals	24
3.6. Equipment.....	24
3.7. Electronic Database Information	26
4. METHODS.....	27
4.1. DNA Isolation.....	27
4.2. Linkage Analyses.....	27
4.3. Candidate Gene Approach	33
4.3.1. Amplification of Candidate Genes by PCR.....	33
4.3.2. Analysis of Amplification Products on Agarose Gels.....	33
4.3.3. Preparation of Denaturing Polyacrylamide Gels.....	34
4.3.4. Denaturing Polyacrylamide Gel Electrophoresis	34
4.3.5. Single Strand Conformational Polymorphism Analysis.....	34
4.3.6. SSCP Gel Preparation.....	35
4.3.7. SSCP Electrophoresis.....	35
4.3.8. Visualization.....	35

4.3.9. DNA Sequence Analysis	36
5. RESULTS	40
5.1. Linkage Analysis	40
5.1.1. Seckel Syndrome	40
5.1.2. Azoospermia	59
5.2. Gene Analysis	63
5.2.1. <i>MNAT1</i>	63
5.2.2. <i>RAD51L1</i>	64
5.2.3. <i>P2RX1</i>	65
5.2.4. <i>NYD-SP20</i>	66
6. DISCUSSION	67
REFERENCES	73

LIST OF FIGURES

Figure 1.1. Pedigree and CFTR genotypes of Azoospermia family	13
Figure 5.1. Linkage results of family 15 at SCKL2	45
Figure 5.2. Silver-stained gel showing the genotypes of some Seckel family members for marker D18S866	43
Figure 5.3. Silver-stained gel showing the genotypes of some Seckel family members for marker D18S877	44
Figure 5.4. Silver-stained gel showing the genotypes of some Seckel family members for marker D18S965	44
Figure 5.5. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S139	46
Figure 5.6. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S586	46
Figure 5.7. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S119	47
Figure 5.8. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S1429	47
Figure 5.9. Silver-stained gel showing the genotypes of some Seckel family members for marker D12S375	57
Figure 5.10. Silver-stained gel showing the genotypes of some Seckel family members for marker D12S1291	57

Figure 5.11. Silver-stained gel showing the genotypes of some Seckel family members for marker D8S1460	57
Figure 5.12. Silver-stained gel showing the genotypes of some Seckel family members for marker D8S1460	58
Figure 5.13. Silver-stained gel showing the genotypes of some Seckel family members for marker D17S1293	58
Figure 5.14. Results of seventeen-point lod scores at chromosome 17p13 for azoospermia using GENEHUNTER	62
Figure 5.15. SSCP gel showing 77205A/G in <i>MNAT1</i>	63
Figure 5.16. SSCP gel showing <i>RAD51L1</i> exon 10	64
Figure 5.17. SSCP gel showing <i>RAD51L1</i> exon 4 and exon 6	64
Figure 5.18. SSCP gel showing <i>RAD51L1</i> exon 5 and exon 2	65
Figure 5.19. SSCP gel showing <i>P2RX1</i> exon 1	65
Figure 5.20. SSCP gel showing <i>P2RX1</i> exon 4 and exon 5	66

LIST OF TABLES

Table 1.1.	Candidate DNA repair genes, their loci, and markers used for linkage analysis for Seckel syndrome	8
Table 1.2.	Seckel syndrome families.....	12
Table 3.1.	Primer sequences for <i>MNAT1</i> amplification	21
Table 3.2.	Primer sequences for <i>P2RX1</i> amplification	22
Table 3.3.	Primer sequences for <i>RAD51L1</i> amplification.....	23
Table 3.4.	Primer sequences for <i>NYD-SP20</i> amplification	23
Table 4.1.	Properties of markers used for homozygosity mapping at gene locus SCKL2.....	29
Table 4.2.	Properties of markers used for homozygosity mapping at gene locus SCKL3.....	30
Table 4.3.	Properties of markers used for homozygosity mapping at candidate DNA repair gene loci	31
Table 4.4.	Properties of markers used for homozygosity mapping at the azoospermia locus	32
Table 4.5.	Sequences of amplified regions of designed markers	32
Table 4.6.	PCR conditions to amplify <i>MNAT1</i>	37
Table 4.7.	PCR conditions to amplify <i>RAD51L1</i>	37

Table 4.8.	PCR conditions to amplify <i>P2RX1</i>	38
Table 4.9.	PCR conditions to amplify <i>NYD-SP20</i>	39
Table 5.1.	Map distances for markers and haplotypes for 15 families at SCKL2. Data from Kılınç [43] are in italics. Noninformative markers are underlined	42
Table 5.2.	Map distances for markers and haplotypes for 15 families at 14q. Data from Kılınç [43] are in italics. Shared genotypes with the unaffected sibs are underlined	48
Table 5.3.	Linkage results of Seckel families around <i>DNAPK</i>	52
Table 5.4.	Linkage results of Seckel families around <i>LIG4</i>	52
Table 5.5.	Linkage results of Seckel families around <i>RAD17</i> and <i>XRCC4</i>	53
Table 5.6.	Linkage results of Seckel families around <i>RAD51L3</i> , <i>LIG3</i> and <i>RAD52B</i>	53
Table 5.7.	Linkage results of Seckel families around <i>RAD9</i>	54
Table 5.8.	Linkage results of Seckel families around <i>RAD52</i>	54
Table 5.9.	Linkage results of Seckel families around <i>XRCC2</i>	54
Table 5.10.	Linkage results of Seckel families around <i>RAD18</i>	54
Table 5.11.	Linkage results of Seckel families around <i>RAD51C</i>	55
Table 5.12.	Linkage results of Seckel families around <i>XRCC3</i>	55

Table 5.13. Linkage results of Seckel families around <i>CHK2</i>	55
Table 5.14. Linkage results around <i>CDK2</i> , <i>CDK4</i> and <i>MDM2</i> for Seckel families whose patients which were homozygous for D12S375	56
Table 5.15. Linkage results around <i>CHK1</i> whose patients were homozygous for D11S4464	56
Table 5.16. Haplotypes of azoospermia family and other azoospermia patients	60
Table 5.17. Two-point lod scores between azoospermia and 17 markers at chromosome 17	61

LIST OF ABBREVIATIONS

A	Adenine
ATM	Ataxia-telangiectesia mutated
ATR	Ataxia-telangiectesia and Rad3 related protein
AZF	Azoospermia factor
APS	Ammonium peroxodisulphate
BRCA1	Breast cancer gene 1
CAK	CDK-activating kinase
CDK	Cyclin dependent kinase
cen	Centromere
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane regulator
CHK1	Cell cycle checkpoint kinase 1
CHK2	Cell cycle checkpoint kinase 2
cM	Centimorgan
dNTP	Deoxynucleoside triphosphate
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetate
G	Guanine
GDI1	GDP dissociation inhibitor 1
GTP	Guanine trinucleotide phosphate
G2	Gap 2
HDAC	Histone deacetylase
H2AX	H2A histone family, member X
IR	Ionizing radiation
Lod	Log of odds
NBS	Nijmegen breakage syndrome
NBS1	Nibrin
NER	Nucleotide excision repair
MDE	Mutation detection enhancement
MDM2	Mouse double minute 2

min	Minute
MNAT1	Menage a trios 1
MRE11	Meitotic recombination 11 homolog
mRNA	Messenger ribonucleic acid
PCR	Polymerase chain reaction
P2RX1	Purinergetic receptor 2X, ligand gated ion channel 1
p53	Tumor protein p53
RAB31	Ras-associated protein
RAD17	RAD17 homolog
RAD51	RAD51 homolog protein
RAD51C	RAD51 homolog C
RAD51L1	RAD51L1-like 1
RAD51L3	RAD51L3-like 3
RBBP1	Retinoblastoma binding protein 1
RBBP8	Retinoblastoma binding protein 8
rpm	Revolution per minute
SCKL	Seckel syndrome
SCKL1	Seckel syndrome 1
SCKL2	Seckel syndrome 2
SNP	Single nucleotide polymorphism
SSCP	Single strand conformation polymorphism
SNP	Single nucleotide polymorphism
SSCP	Single strand conformation polymorphism
T	Thymine
Tel	Telomere
TEMED	N, N, N, N'-Tetramethylethylenediamine
U	Unit
UTR	Untranslated region
UV	Ultraviolet
W	Watt
XRCC2	X-ray repair cross complementing protein 2
XRCC3	X-ray repair cross complementing protein 3
XRCC4	X-ray repair cross complementing protein 4

1. INTRODUCTION

In this study, genetic analyses were performed with the purpose of refining known loci and identifying genes responsible for two autosomal recessive disorders: Seckel syndrome and azoospermia. Also, additional gene loci were investigated for Seckel syndrome.

1.1. Seckel Syndrome

Seckel syndrome (SCKL) is a rare autosomal recessive disorder with heterogeneous phenotype. Basic features are growth retardation, microcephaly with mental retardation, and a characteristic bird-headed profile (SCKL [MIM 210600]). In addition to dwarfism of 'low birth weight' type, the common features are small head, large eyes, beak-like protrusion of the nose, narrow face, and receding lower jaw. Typical very small brain size results in a relatively lesser degree mental retardation than expected for the reduced size. Multiple occurrences in the same sibship, a high frequency of parental consanguinity, occurrence in both sexes, and normal parents clearly indicate autosomal recessive inheritance.

A number of Seckel-like syndromes also have been defined, such as microcephalic osteodysplastic primordial dwarfism types I-III [1, 2, 3, 4] and microcephalic osteodysplastic dysplasia [5], which can be differentiated from one other by radiographical and clinical analyses. Many of the so-called Seckel cases may in fact be these Seckel-like disorders.

Seckel syndrome exhibits very high clinical heterogeneity. More than 100 cases of SCKL have been diagnosed. Although they all share several common clinical findings, some of the features are present only in a few cases, and some are seen only in single cases, such as atrioventricular canal defect, nonclassical congenital adrenal hyperplasia, cleft lip and palate, dislocation of lenses, and acute myeloid leukemia [6]. Blasi *et al.* previously suggested that Seckel cases were over diagnosed, and many of the reported cases were not truly Seckel [7]. The SCKL cases reported share some of the common

features first reported by H. P. G. Seckel in 1960 but exhibit high phenotypic heterogeneity, supporting the hypothesis that genetic defects in different genes can lead to SCKL.

1.1.1. The Seckel Loci

Three loci have been identified for Seckel syndrome. The first is SCKL1, which is at chromosome 3q22.1-q24, and the gene responsible has been identified [8]. The second locus SCKL2 is at chromosome 18p11.31-q11.2 [9]. The last one, SCKL3, is at chromosome 14q23 [10].

1.1.1.1. SCKL1. Goodship *et al.* carried out autozygosity mapping in two consanguineous Seckel families from the same village in Pakistan but with no known relationship to each other [8]. After a genomic screen in these families, they found that the disorder segregated with markers between D3S1316 and D3S3710, which map to chromosome 3q22.1-q24 (12cM). The patients were homozygous for the same alleles, indicating a common ancestor. The gene *ATR*, residing in this chromosomal region, was a strong candidate. *ATR* is a gene related to *ATM*, which is known to cause ataxia-telangiectasia when defective. *ATM* has several roles including signaling of DNA damage and control of cell cycle progression. As Seckel syndrome shares many clinical features with other DNA repair syndromes, it is highly probable that genes functioning in DNA repair mechanisms are responsible for Seckel syndrome.

Other previous findings also supported the hypothesis of Seckel syndrome being a DNA repair defect. Syrrou *et al.* described increased frequency of chromosome abnormalities in response to mitomycin C and increased frequency of sister-chromatid exchange in three siblings with Seckel [11]. Woods *et al.* also reported a newborn with clinical diagnosis of Seckel syndrome in whom chromosome breakage following mitomycin treatment was observed [12]. In this child, some symptoms in common with Nijmegen breakage syndrome (NBS) were reported. The defective gene in NBS encodes *p95*, which is a member of the MRE11/RAD50 double-strand break-repair complex [13].

O'Driscoll *et al.* [14] found that a synonymous mutation disrupting proper splicing of *ATR* gene (ataxia-telangiectasia and RAD3-related protein) was responsible for the disease in the Pakistani families studied by Goodship *et al.* [8]. The patients were homozygous for this mutation. They used a fibroblast cell line (F02-98) from an affected individual and showed that the cells displayed a defective DNA damage response caused by *ATR* impairment. This synonymous mutation (2101 A→G) results in reduced but residual levels of normal transcript and protein. *ATR* is a phosphatidylinositol 3-kinase-like kinase, and like *ATM*, functions for the most part in response to DNA double-strand breaks caused by ionizing radiation. Its major role is induction of DNA damage response to other forms of damage, such as stalled-replication forks, which are induced by ultraviolet radiation. After DNA damage, *ATR* phosphorylates substrates including p53, Rad17, Nbs1 and H2AX, which are also common substrates for *ATM*. The analysis of *ATR* function in F02-98 cells revealed improper response to ultraviolet radiation exposure. Phosphorylation of the substrates of *ATR* in F02-98 cells was impaired, and the level of *ATR* expression was reduced. In contrast, response of these cells to ionizing radiation, which is mediated by *ATM*, is normal. These results revealed that defective *ATR* impaired the response to DNA damage [14].

1.1.1.2. SCKL2. Borglum *et al.* identified the Seckel locus at 18p11.31-q11.2 in a consanguineous family of Iraqi descent [9]. Identification of this new locus was the first evidence of genetic heterogeneity in Seckel syndrome, which was foreseen by clinical heterogeneity. This new locus was nearly 30 cM wide and included several candidate genes. Microsatellite markers D18S78 and D18S866 flanked the gene region, which was nearly 15 Mb.

Seckel syndrome shares several phenotypic characteristics with disorders of DNA repair. After identification of *ATR* as a defective gene in a Seckel kindred, genes that have roles in DNA repair were regarded as strong candidates for the disorder. At the locus SCKL2, there are two strong candidate genes: *Retinoblastoma Binding Protein 8 (RBBP8)* and *Ras-associated protein (RAB31)*.

1.1.1.3. SCKL3. SCKL3, which maps to 14q23, was identified by Kılınç *et al.* in our laboratory [10]. Identification of this novel locus further confirmed the genetic

heterogeneity of Seckel syndrome. Patients in one of the families were homozygous for a 12,672 Mb region at this locus. Initial analyses revealed a maximum lod score around marker D14S592. Patients in two other families were homozygous for this and another one for the subsequent marker D14S586. The defective gene was proposed to reside around D14S592. The strongest candidate was *Menage A Trois 1*, (*MNAT1*) having a role in DNA repair.

1.1.2. Candidate Genes at SCKL2

After analysis of databases for the genes at SCKL2 locus, two genes were assessed as strong candidates.

1.1.2.1. *RBBP8*. *Retinoblastoma binding protein 8* is a ubiquitously expressed nuclear protein that directly binds to retinoblastoma protein, which regulates cell proliferation. *RBBP8* (*CTIP*) is associated also with *BRCA1* and thought to modulate the functions of *BRCA1* in transcriptional regulation, DNA repair, and/or cell cycle checkpoint control. This gene encodes two different protein isoforms [15].

1.1.2.2. *RAB31*. This gene encodes the small GTP-binding rab protein. The expression pattern of the protein interestingly correlates with the affected organs in Seckel syndrome. *Rab31* protein is involved in the vesicular transport of molecules between the cellular organelles. *GDI1* protein, which modulates rab family proteins, was found to be defective in a number of families with X-linked mental retardation [16].

1.1.3. Candidate Genes at SCKL3

The database information on genes residing at SCKL3 revealed three strong candidates.

1.1.3.1. *MNAT1*. *Menage A Trois* (*MNAT1*), which contains a ring finger domain at its N terminus, is a subunit of a trimeric protein complex which is called CDK-activating kinase (*CAK*). This complex phosphorylates and hence activates cyclin-dependent kinases (*CDKs*), which play crucial roles in cell cycle control of eukaryotic cells (OMIM 602659).

MNAT1 is a must primarily for the stimulation of kinase activity of CAK and secondarily for the assembly of this trimeric complex, which is also required for basal transcription and DNA repair. CAK also forms a part of a multisubunit transcription factor TFIIH. This complex contains additionally XPD (XERCC2) and XPB (XRCC3), both helicases involved in nucleotide excision repair (NER). This interaction underlines the probable role of MNAT1 in DNA repair [17]. The gene is located at 59,1 – 59,4 Mb region on chromosome 14. It has eight exons and codes for a 309 amino-acid peptide.

1.1.3.2. *RAD51L1*. *Escherichia coli* RecA protein provides an enzymatic function for recognition of DNA sequence homology and DNA strand exchange that are essential steps in the molecular systems necessary for genetic recombination and recombinational repair [18]. RAD51 protein of *Saccharomyces cerevisiae*, which is homologous to Rec A, operates in both mitotic and meiotic homologous recombination as well as in double-strand break repair (OMIM 602948). *RAD51L1* is a member of the RAD51 family whose members are evolutionarily conserved genes essential for DNA repair by homologous recombination. Albala *et al.* identified a human homolog of *RAD51* and designated it *RAD51L1* [19]. It was found that RAD51L1, RAD51C and XRCC2, all DNA repair proteins, were immunoprecipitated with antibodies raised against RAD51L3 [20]. *RAD51L1* gene is located at 14q23-24 [21] and like ATR, which is mutated in SCKL1, shown to be activated in response of UV-induced DNA damage [22]. *RAD51L1* gene spans the genomic region 66,2 – 66,9 Mb on chromosome 14 and codes for 11 exons. Two isoforms differ in their last exons: one includes exon 10 and the other exon 11.

1.1.3.3. *RBBP1*. *Retinoblastoma binding protein 1 (RBBP1)* is a ubiquitously expressed nuclear protein. Together with several other proteins, it directly binds retinoblastoma protein, which attenuates transcription. By binding retinoblastoma protein, RBBP1 serves as a bridging molecule to recruit HDACs and, furthermore, it provides a second HDAC-independent repression function [21].

1.1.4. ATR Related Proteins

Since *ATR* is one of the genes responsible for Seckel syndrome, proteins that directly interact with ATR are good candidates for the disease.

ATR-interacting and other candidate DNA repair genes analyzed for linkage in this study and their locations are presented in Table 1.1.

1.1.4.1. *CHK1*. This protein is encoded by the gene *Cell Cycle Checkpoint Kinase 1* (*CHK1*), an ATR substrate in the G2-M checkpoint pathway [23, 24]. *CHK1* protein complex binds both single-stranded DNA and DNA ends and has a role in the repair of double-strand breaks. In addition, *CHK1* stimulates DNAPK-dependent end-joining reactions in vitro [25]. *CHK1* mediates both S and G2 checkpoints and is thus a ubiquitous cell cycle checkpoint mediator [26]. *CHK1* interconnects pathways that directly regulate DNA repair and checkpoint control [27].

1.1.4.2. *CHK2*. *Checkpoint kinase 2* (*CHK2*) is a protein kinase activated in response to DNA damage and consequential for cell cycle arrest. This protein phosphorylates CDC25C on the same residue phosphorylated by *CHK1* [28]. *CHK2* gene that encodes *CHK2* protein was proposed to act as a tumour suppressor in some of the patients with the cancer predisposing Li-Fraumeni syndrome, which is also a disorder caused by impaired DNA damage response [27]. Phosphorylation of p53 on serine-20 by *CHK2* caused dissociation of preformed complexes of p53 and MDM2 in vitro. MDM2 functions to target p53 for degradation. Therefore, in response to DNA damage, *CHK2* protein stabilizes p53 and leads to cell cycle arrest [29].

1.1.4.3. *RAD17*. Bao *et al.* uncovered that *RAD17*, which is highly similar to the gene product of *Saccharomyces pombe* *rad17*, a checkpoint protein vital for both cell cycle arrest and DNA repair in response to DNA damage, functioned in a regulatory network with checkpoint kinases ATM and ATR [30]. Exposure of human cells to genotoxic agents resulted in phosphorylation of *RAD17* at both serine-635 and serine-645 in an ATM/ATR dependent manner. In addition, disruption of these phosphorylation sites abrogated DNA damage induced G2 checkpoint and caused sensitization of human fibroblasts to genotoxic stress [30]. Therefore, ATR/ATM-dependent phosphorylation of *RAD17* is a vital early event through checkpoint signaling in DNA-damaged cells.

1.1.4.4. *RAD9*. This gene product is highly similar to *S. pombe* *rad9*, a checkpoint protein important in cell cycle arrest and DNA damage repair. Human homolog of this protein

RAD9 was found in a complex with RAD1 and HUS1 [31]. This complex rapidly migrates to sites of DNA damage and acts as checkpoint sensor during S phase [32]. Volkmer and Karnitz [31] proposed that these proteins play a central role in a DNA-damage responsive protein complex in human cells by recruiting and organizing other DNA repair proteins [32, 33]. RAD9 is phosphorylated in an ATR/ATM dependent manner, similar to RAD17.

1.1.5. Other Candidate Repair Genes

1.1.5.1. *LIG4*. *DNA ligase IV (LIG4)* functions in the ligation step in non-homologous DNA end joining and in V(D)J recombination. Mutations in the gene are associated with Lig4 syndrome [34, 35]. The phenotypes of patients with Lig4 syndrome closely resemble that of patients with Nijmegen breakage syndrome. Some of the identified mutations in *LIG4* gene reside in ligase domain and others in the XRCC4 interacting domain. XRCC4 also functions in the non-homologous end-joining repair. Many of the symptoms of Lig4 syndrome closely resemble those of Seckel condition, but, in addition, pancytopenia is commonly seen [35]. Therefore, it would be of great importance to screen Seckel patients for lig4 gene locus, because some of the patients with so-called Seckel syndrome may not be true Seckel, as diagnosis with absolute certainty is not possible.

1.1.5.2. *DNAPK*. *DNA dependent protein kinase (DNAPK)* is also implicated in DNA repair. On its own, DNAPK is inactive, and to be active, it must interact with KU protein, which is encoded by gene G22P1. KU protein specifically binds DNA double-strand breaks to repair this type of damage. DNAPK functions in the formation of a synaptic complex containing broken ends of DNA, and then catalyzes subsequent steps required for further processing and ligation of the DNA ends, which may explain the activation of DNAPK only at those sites. Cells that lack DNAPK are hypersensitive to apoptosis caused by ionizing radiation, which induces double-strand breaks [33].

1.1.5.3. Other DNA Repair Proteins. The other genes analyzed for linkage in the context of this study that have roles in DNA repair were as follows: *RAD18*, *XRCC4*, *XRCC3*, *XRCC2*, *RAD52*, *CDK2*, *CDK4*, *MDM2*, *LIG4*, *RAD51*, *RAD51L3*, *RAD51C*, *LIG3* and *RAD52B*.

Table 1.1. Candidate DNA repair genes, their loci, and markers used for linkage analysis for Seckel syndrome

Gene	Locus	Markers
<i>RAD18</i>	3p25-p24	D3S4545
<i>RAD17</i>	5q13	D5S2500, D5S1725
<i>XRCC4</i>	5q13-q14	D5S1725
<i>XRCC2</i>	7q36.1	D7S1823
<i>DNAPK</i>	8q11	D8S1460, D8S1716, D8S1110
<i>RAD9</i>	11q13.1-q13.2	D11S2371
<i>CHK1</i>	11q24	D11S4464, D11S4158, D11S968
<i>RAD52</i>	12p13-p12.2	D12S372
<i>CDK2</i>	12q13	D12S601, D12S1291, D12S375,
<i>CDK4</i>	12q14	D12S601, D12S1291, D12S375,
<i>MDM2</i>	12q14.3-q15	D12S601, D12S1291, D12S375 D12S1043 D12S1025, D12S828, D12S376
<i>LIG4</i>	13q33-q34	D17S1396
<i>XRCC3</i>	14q32.3	D14S1426
<i>RAD51</i>	15q15.1	GATA50C03
<i>RAD51L3</i>	17q11	D17S1293
<i>LIG3</i>	17q11.2-q12	D17S1293
<i>RAD52B</i>	17q11.2	D17S1293
<i>RAD51C</i>	17q22-q23	D17S1290
<i>CHK2</i>	22q11	D22S689

1.1.6. DNA Repair Syndromes

Genetic defects that disrupt normal DNA damage response pathways lead to a wide variety of syndromes, which may exhibit a combination of various degrees of growth and mental retardation, tissue degeneration, signs of premature ageing, defects in hematopoietic or immune systems, neurological abnormalities, sensitivity towards mutagens, and cancer predisposition. Besides Seckel syndrome (SCKL), the major

genomic instability syndromes are ataxia-telangiectasia (AT), AT-like disorder, Nijmegen breakage syndrome, xeroderma pigmentosum (XP), Cockayne syndrome, trichothiodystrophy (TTD), Bloom's syndrome (BS), Rothmund-Thompson syndrome (RTS), Werner's syndrome (WS), and Fanconi's anemia (FA).

1.2. Azoospermia

Cystic fibrosis (CF) was the first genetic disorder to be mapped by positional cloning. The gene responsible for this autosomal recessive disease is the *cystic fibrosis transmembrane conductance regulator* (CFTR). The protein encoded by CFTR gene functions as a chloride channel and controls the regulation of other transport pathways. Mutations in the CFTR gene have been found to also cause isolated congenital bilateral aplasia of the vas deferens (CBAVD).

The great majority of men with cystic fibrosis (CF) have associated CBAVD, the absence of the anatomic ducts through which sperm passes from the testes to the urethra. No sperm is found in the semen, a defect referred to as azoospermia. This defect in male genital system is not a developmental anomaly but a degenerative alteration due to blockage similar to that in the pancreas and salivary glands in cystic fibrosis. This physiological defect causes infertility in men and may also be seen in isolation, with no CF background. As this condition was commonly seen in CF cases, patients with CBAVD were suspected to carry CF mutations. Indeed, many CBAVD cases were shown to carry CFTR mutations on one or both alleles [36].

Rave-Harel *et al.* tested the hypothesis of association of CBAVD and CF. According to his hypothesis, 2 siblings with CBAVD would be expected to carry the same CFTR genotype. However, in one of the eleven families studied, two brothers with CBAVD did not have the same CFTR genotypes, and one fertile brother has same CFTR alleles as one of the brothers with CBAVD. In one other family one fertile brother had the same CFTR genotype as a sterile [36]. Also, in another family, two brothers with the same CFTR genotype displayed different phenotypes, one being fertile and the other sterile [37]. These findings suggested that additional genes had roles in the development of azoospermia. The

5T allele of the poly-thymine tract in intron 8 of *CFTR*, which causes reduced levels of the normal *CFTR* mRNA, is also involved in CBAVD [38].

Cytogenetic analyses in infertile men suggested that the locus Yq11.23 controlled spermatogenesis and was referred to as azoospermia factor (AZF) [39, 40]. Most of the patients with azoospermia are chromosomally normal but may have microdeletions in AZF sequences. These microdeletions probably affect the functional DNA structure of the AZF gene [40]. So, screening of azoospermia patients for Y chromosomal deletions at those AZF sequences may uncover the underlying cause of the condition in some cases.

1.2.1. Candidate Genes for Azoospermia

Two genes were assessed as strong candidates at the azoospermia locus.

1.2.1.1. *P2RX1*. *Purinergic receptor P2X, ligand-gated ion channel, 1 (P2RX1)* protein belongs to the family of purinoceptors for ATP. It functions as a ligand-gated ion channel with relatively high calcium permeability. It binds ATP and mediates synaptic transmission between neurons and smooth muscle. It is responsible for sympathetic vasoconstriction in small arterioles and vas deferens. In a mouse model, it was suggested to be essential for normal male reproductive function [41].

1.2.1.2. *NYD-SP20*. This gene encodes a testis-development related protein. It is differentially expressed in testis. It shares homology with NYD-SP16, which had weak or no expression in patients with arrest at the spermatid stage and but high in patients that had spermatogenic cells of every stage [42].

1.3. Families

In this study, fifteen Seckel families were analyzed. Nine of the fifteen families were small, with only one affected child or one affected and one healthy children. Despite the small size, consanguineous families contribute to the cumulative lod score when the patients exhibit homozygosity at the locus. DNA samples from only two generations were available for all families. Previously, a genome-wide search was performed in our

laboratory in three of the large families [13]. One of them linked to SCKL1 and the other to SCKL3. No significant linkage was determined in the third family, but homozygosity for two adjacent markers at SCKL3 was found. Autozygosity mapping was required for the other families for testing linkage in the identified loci. The families that were not linked to known loci could be investigated further to determine other loci/genes implicated in Seckel. However, most of those families were too small to yield significant lod scores. Nevertheless, all families were analyzed to narrow down the known loci and to investigate linkage to selected candidate gene loci, as the accumulative data would be significant. In addition, the patients were screened for defects in candidate genes. The families used in Seckel studies are presented in Table 1.2. In the tables and figures through text the individuals are designated relative to their relation to the patients (P: patient, F: father of the patient, M: mother of the patient, S: sibling of the patient).

In the azospermia family, there were 3 affected and 1 unaffected boy. One sister was also included to help deduce father's genotype, as his DNA was not available. Because of the consanguinity, the family provided excellent genetic material to search for a putative second locus. It was a family member who had applied to us for a genetic study. The large family used in azoospermia studies is presented in Figure 1.1. In this figure, the letter N designates normal wild type CFTR allele.

Initial genetic studies had been performed in our laboratory for most of the Seckel families and the azoospermia family. Kılınç [43] analyzed Seckel syndrome patients in 13 Anatolian families. A genome-wide scan was performed in families 1, 2 and 3, which were selected because of their large sizes. Before it was completed, Seckel locus at 3q22.1-24 (SCKL1) was reported by Goodship *et al.* [8]. Kılınç [43] determined linkage in family 1 to that locus and excluded the remaining 12 families. The maximum region of homozygosity spanned a 9.582 Mb region delineated by markers D3S3528 and D3S1608, and the minimal critical region was 5.915 Mb between D3S1316 and D3S2394.

Table 1.2. Seckel syndrome families

Family 1	F1	M1	P1.1	P1.2	S1.1	S1.2	S1.3
	Father	Mother	Patient	Patient	Healthy	Healthy	Healthy
Family 2	F2	M2	P2.1	P2.2	S2.1	S2.2	S2.3
	Father	Mother	Patient	Patient	Healthy	Healthy	Healthy
Family 3	F3	M3	P3.1	P3.2	S3		
	Father	Mother	Patient	Patient	Healthy		
Family 4	F4	M4	P4.1	P4.2			
	Father	Mother	Patient	Patient			
Family 5	F5	M5	P5	S5.1	S5.2	S5.3	S5.4
	Father	Mother	Patient	Healthy	Healthy	Healthy	Healthy
Family 6	F6	M6	P6				
	Father	Mother	Patient				
Family 7	F7	M7	P7				
	Father	Mother	Patient				
Family 8	F8	M8	P8.1	P8.2	S8		
	Father	Mother	Patient	Patient	Healthy		
Family 9	F9	M9	P9				
	Father	Mother	Patient				
Family 10	F10	M10	P10	S10			
	Father	Mother	Patient	Healthy			
Family 11	F11	M11	P11				
	Father	Mother	Patient				
Family 12	F12	M12	P12				
	Father	Mother	Patient				
Family 13	F13	M13	P13				
	Father	Mother	Patient				
Family 14	F14	M14	P14				
	Father	Mother	Patient				
Family 15	F15	M15	P15	S15			
	Father	Mother	Patient	Healthy			

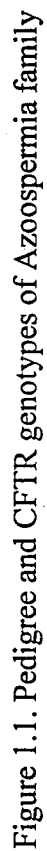


Figure 1.1. Pedigree and CFTR genotypes of Azoospermia family

After continuing the genome-wide search, Kılınç [43] determined linkage to chromosome 14 in families 1, 2, and 3. In family 3, the region was in a minimal 9.536 Mb critical region between markers D14S586 and D14S588 and in maximal region between D14S601 and D14S77, spanning 14.087 Mb. However, in this region, patients in families 1 and 3 were homozygous only for two markers (D14S586 and D14S592) resulting in a minimal critical region of 0.594 Mb. These markers were flanked by D14S601 and D14S997, which were 3.217 Mb apart. The locus was designated as SCKL3 by Kılınç [43]. Interestingly, family 3 linked to both SCKL1 and SCKL3. Whether this observation has a molecular basis could be uncovered only after identification of both genes and screening them for mutations in the patients. Also, seven families were roughly analyzed at SCKL2 by Uğur [44].

In the azoospermia family, as the father was reportedly CF, *CFTR* gene had been analyzed. Two of the affected sibs but not the third one carried the novel D891G mutation. This finding suggested that a gene at another locus was responsible for the disorder. Also, the family had been analyzed for poly-dT tract in intron 8 of *CFTR* gene but found to carry the common 7T allele rather than 5T, which is associated with azoospermia [43]. After a genome-wide scan in the azoospermia family, a locus at chromosome 17 had been identified [43].

1.4. Strategy of the Study

Three loci have been reported for Seckel syndrome so far, and it is possible that others exist, as well. Only the disease gene at one locus (SCKL1) has been elucidated; the other genes implicated in Seckel syndrome have not yet been identified. In order to identify the Seckel genes at SCKL2 and SCKL3, narrowing down the regions would facilitate a better assessment of candidate genes. Such genes at each locus can be assessed using the information in databases and the strongest candidates selected for analysis.

In the first part of the study, the purpose was to analyze Seckel families to test linkage to three known disease loci. Furthermore, the study aimed at narrowing down those loci in order to better assess the candidate genes. In addition, as Seckel syndrome is a genetically heterogeneous disease, some other candidate loci were analyzed. A set of

candidate genes that have roles in DNA repair was selected and Seckel families tested for linkage by homozygosity mapping using nearby microsatellite markers in order to determine linkage. Azoospermia was the second disorder studied. The locus that was identified by Kılınç [43] was refined using additional markers in order to better assess the strongest candidate genes.

In the second part of the study, by utilizing the information in genome databases, two genes that have roles in DNA repair were determined and assessed as the strongest candidate genes for SCKL3, and two genes that have roles in testis development were determined and assessed as strong candidate genes, at relevant loci. The four candidate genes were *MNAT1* and *RAD51L1* for Seckel syndrome, and *P2RX1* and *NYD-SP20* for azoospermia. They were all analyzed for putative mutations in the patients in families that were indicative of linkage to the gene loci.

2. PURPOSE

A high incidence of consanguineous marriages in Turkey, which leads to a high incidence of very rare autosomal recessive disorders, is invaluable for human genetic studies. A large number of novel diseases have been identified in our population. As most of those disorders are very rare, our studies would not have direct therapeutic importance but will contribute to the understanding of molecular pathways and cellular mechanisms.

In the first part of the study, homozygosity mapping was performed in order to find out any significant linkage to known genetic loci, narrow down the known loci, and investigate significant linkage to ATR-interacting and other DNA repair genes using markers within the gene loci.

In the second part of the study, the strongest candidate genes at the loci of interest were selected by utilizing the information in genome databases. Candidate gene approach was used for the purpose of identifying any causative mutations in four candidate genes: *MNAT1* and *RAD51L1* for Seckel syndrome and *P2RX1* and *NYD-SP20* for azoospermia.

3. MATERIALS

3.1. Blood Samples

Peripheral blood samples from families with Seckel/Seckel-like syndromes used in this project were provided by Prof. S. Balcı at Hacettepe University in Ankara, Prof. B. Tüysüz at Cerrahpaşa Medical School in İstanbul, Dr. M. Aynacı at Karadeniz Technical University in Trabzon, and Dr. B. Eroğlu at Şişli Etfal Hospital in İstanbul.

Peripheral blood samples from patients with azoospermia and their relatives used in this study were supplied by Dr. A. Tatar.

3.2. Chemicals

All chemicals used in this study were purchased from Sigma (USA) or MERCK (Germany) unless stated otherwise in the text.

3.3. Buffers and Solutions

3.3.1. DNA Extraction from Blood Samples

Cell Lysis Buffer : 155 mM NH_4Cl , 10 mM KHCO_3
0.1 mM Na_2EDTA (pH 7.4)

Nucleus Lysis Buffer : 400 mM NaCl , 2 mM Na_2EDTA ,
10 mM Tris-HCl (pH 8.2)

Proteinase K : 20 mg/ml Proteinase K in dH_2O
(Promega, USA)

Sodiumdodecylsulfate (SDS) : 10 per cent (w/v) SDS in dH_2O

Ammonium Acetate	:	7.5 M CH ₃ COONH ₄ in dH ₂ O
Ethanol	:	Absolute Ethanol (Riedel de-Haen, Germany)
TE buffer	:	1 mM EDTA, 20 mM Tris-HCl (pH 8.0)

3.3.2. Polymerase Chain Reaction (PCR)

10 X PCR Buffer	:	15 mM MgCl ₂ , 500 mM KCl 100 mM Tris-HCl (pH 8.3) (Roche, Germany)
10 X PCR Buffer (Mg free)	:	500 mM KCl 100 mM Tris-HCl (pH 8.3) (Roche, Germany)
MgCl ₂	:	25 mM MgCl ₂ (Promega, USA)
dNTP	:	12.5 mM each of dATP, dCTP, dGTP and dTTP in dH ₂ O
DNA Polymerase	:	Promega (USA) and Roche (Germany)

3.3.3. Agarose Gel Electrophoresis

Agarose	:	2 per cent Agarose in 0.5 X TBE buffer
10 X TBE Buffer	:	20 mM EDTA, 0.89 M Boric Acid, 0.89 M Trizma base (pH 8.3)
10 X Loading Buffer	:	2.5 mg/ml Bromophenol Blue, 1 per cent SDS in Glycerol
Ethidium Bromide	:	20 mg/ml in dH ₂ O

3.3.4. Denaturing Polyacrylamide Gel Electrophoresis (PAGE)

40 per cent Acrylamide (Stock)	:	40 per cent Acrylamide-bisacrylamide (19:1) in dH ₂ O
8 per cent Instagel	:	8 per cent Acrylamide-bisacrylamide (19:1), 8.3 M urea in 1 X TBE buffer
APS	:	10 per cent Ammonium Peroxidisulfate
TEMED	:	N,N,N,N-tetramethylethylenediamine
10X Loading Buffer	:	95 per cent Formamide, 20 mM EDTA, 0.05 per cent Xylene Cyanol, 0.05 per cent Bromophenol Blue

3.3.5. Single Strand Conformational Polymorphism (SSCP) Gel Electrophoresis

40 per cent Acrylamide (Stock)	:	40 per cent Acrylamide-bisacrylamide (37.5:1 or 50:1) in dH ₂ O
8 per cent Acrylamide (Non-denaturing)	:	8 per cent Acrylamide-bisacrylamide (37.5:1 or 50:1) in 0.6 X TBE buffer
Glycerol	:	5 per cent Glycerol in SSCP gel solution

3.3.6. Silver Staining

Staining Buffer (Buffer B)	:	0.1 per cent AgNO ₃ (w/v) in dH ₂ O
----------------------------	---	---

Developing Buffer (Buffer C) : 1.5 per cent NaOH,
0.01 per cent NaBH₄,
0.015 per cent Formaldehyde in dH₂O

Stop Buffer (Buffer D) : 0.75 per cent NaCO₃ in dH₂O

Storage Buffer (Buffer E) : 5 per cent Glycerol in dH₂O

3.4. Oligonucleotide Primers

A total of 130 oligonucleotide primer pairs were used in this study. These were either microsatellite markers for linkage analysis or primers for amplification of candidate genes.

Fine mapping was performed with a total of 88 microsatellite markers, which were purchased from either Research Genetics (USA) or Iontek (Turkey). Eighty-four of those were available at NCBI web site (National Center for Biotechnology Information, <http://www.ncbi.nih.gov/unists>). The remaining four were designed in the context of the study, as no available markers were reported at those regions. For designing markers, regions with repeated sequences were determined by using the software Tandem Repeats Finder Version 3.21, which was available at the web site <http://tandem.bu.edu/trf/trf.html>. The strategy for choosing appropriate repeat regions was as follows: Tetra repeats with 15-30 copy numbers were chosen wherever possible, since tetra repeats were in general found to be highly polymorphic. If they were not present, tri or di repeats with copy numbers around 40-80 were selected. Sequences of the designed markers are given in Table 4.5. Resgen markers were purchased already at 20 µM concentration, while markers purchased from Iontek needed to be diluted to 20 µM.

The oligonucleotide primers for amplification of the candidate genes were purchased from Iontek (Turkey). Stock primers were diluted to 10 µM in 60 µl dH₂O for PCR reactions. A total of 42 primer pairs were used to analyze four candidate genes. Primers were designed to amplify flanking intronic regions (25-30 bp) to allow investigation of conserved sequences around splice sites. Primer design was carried out using the software Primer3, which was available at <http://workbench.sdsc.edu>. The sequences of primers for analysis of candidate genes are presented in Table 3.1, 3.2, 3.3, and Table 3.4. The sets of

microsatellite markers used in the context of this study are given in Tables 4.1, 4.2, 4.3, and 4.4. Exons 1 and 4 of *NYD-SP20* were amplified in 2 overlapping fragments each due to the large exon sizes.

Table 3.1. Primer sequences for *MNAT1* amplification

Exon	Primer Name	Primer Sequence (5'→3')
1	MNAT1-1F MNAT1-1R	GTTGGTAGGAACCTGCTTGG CATTTTCCTCCCCAGTGTTG
2	MNAT1-2F MNAT1-2R	TCCCATAGGGGCAAACCTATTC GCTAATCTGTCCACATTTATGAGC
3	MNAT1-3F MNAT1-3R	GGTCAGAACAAGCAAACCAG GCTATCACAACAAAACCTGTTAAG
4	MNAT1-4F MNAT1-4R	CCTAACAAATGAGTCACTGTGG CTGGTTTAAAGGAAGATGGC
5	MNAT1-5F MNAT1-5R	TGATTGCTCTTGGCTCCTAAA AATGCCATCATCCAGCCTA
6	MNAT1-6F MNAT1-6R	GTTTGACATGAAATGCTTATGG CAGACACTGTTCTTAAACATCTGAA
7	MNAT1-7F MNAT1-7R	CATGGAACAGCACAACTAATGG TCCCAACATATCAGCTCCTCAAA
8	MNAT1-8F MNAT1-8R	GGCACTCCATGTACAGTTGGTA CTTTTGCTAGCTCCCTGGTTC

Table 3.2. Primer sequences for *P2RX1* amplification

Exon	Primer Name	Primer Sequence (5'→ 3')
1	P2RX1-1F P2RX1-1R	GAAGCCCCCAGAAGCTCTAC ACGCAAGGGATGTCAGAGTC
2	P2RX1-2F P2RX1-2R	AAGCGTGTGACAGGTCCAG AGAGTCCCACCCAACTCTGA
3	P2RX1-3F P2RX1-3R	CTCTTCTCCCAGCTCCCAG TCATTTTCCCCTGTGAGTCC
4	P2RX1-4F P2RX1-4R	AGAGGGGCTCTTCTCAGGAC GGTGTGGAGAGAGGCAGGT
5	P2RX1-5F P2RX1-5R	TTCTGTCCTTGGCTCCTTTG GGGGTCAGAAAAAGGGGTAA
6	P2RX1-6F P2RX1-6R	GCTGTGCCAAGCTCTGTGT ATTGCCCCCTTTGATTCTTC
7	P2RX1-7F P2RX1-7R	GAAGGAATCAAAGGGGCAAT CGAGGAGACTTGGAAGATG
8	P2RX1-8F P2RX1-8R	CTCCCAAAGTGCTGGGATTA TGGGGACACTTCTCAGACC
9	P2RX1-9F P2RX1-9R	CCTGATGCGCTAAGAACTCC AAACTGCTGGATGAACCCAC
10	P2RX1-10F P2RX1-10R	CATGTAGAGTCCGAGGTGGG GAGGGACGTCTTGCTGCTAC
11	P2RX1-11F P2RX1-11R	CCTCCAGTGAGGTCAGATCC CACCCCAACACAGACACAAG
12	P2RX1-12F P2RX1-12R	CTTGTGTCTGTGTTGGGGTG GGTAGGAGACTTCCTGGGGA

Table 3.3. Primer sequences for *RAD51L1* amplification

Exon	Primer Name	Primer Sequence (5'→ 3')
1	RAD51L1-1F	TGCATATGGAGAACTTGAGGA
	RAD51L1-1R	GCCTATGGAAAAGCACCATC
2	RAD51L1-2F	AAGCAGTATCAATTGAAGGTTTCC
	RAD51L1-2R	GCATGTGCCTCATCCCTAAA
3	RAD51L1-3F	GGGAAAAGGATGGATTAGGG
	RAD51L1-3R	TTTTCCCTAGCTCTGAAGGTT
4	RAD51L1-4F	TTTGCTTTGACTGGCTTGTG
	RAD51L1-4R	CCTATGTGTAAATGTTTCATGCTG
5	RAD51L1-5F	GAGATTCAGCAATGTGGCTTT
	RAD51L1-5R	AAAACATAAACCTTCAACGCA
6	RAD51L1-6F	TTGCTTAGCCTTTGCAATACC
	RAD51L1-6R	GAAACAAAAGAAAGGAAAAGAAAA
7	RAD51L1-7F	GTCTTTCTTCTCCCTTGCCC
	RAD51L1-7R	AGCTCTCAACAGGAGGTGGA
8	RAD51L1-8F	CCATCTCAAGAAAGGGCATC
	RAD51L1-8R	TAATAGCTGCTGCTGGCTCA
9	RAD51L1-9F	TTGTCTCAAAGTGGGAAGGG
	RAD51L1-9R	TACCACTAGCAATGGGCAGC
10	RAD51L1-10F	AATGGGGGCATGAACTTTTT
	RAD51L1-10R	GCAGAGAACGATGGGTGTTT
11	RAD51L1-11F	TTTAGGTTGCTGGTGGGAAG
	RAD51L1-11R	CCCATTATGTCAGCTTCCGT

Table 3.4. Primer sequences for *NYD-SP20* amplification

Exon	Primer Name	Primer Sequence (5'→ 3')
1a	NYD-SP20-1aF	GGATGGTTGTGGAATTTTAGC
	NYD-SP20-1aR	AAACTTTCCAGGCATTTGTCA
1b	NYD-SP20-1bF	AATTAAGCAATAACAGAAAACCTGCT
	NYD-SP20-1bR	TTTTGAGGTGGCACATGTTT
2	NYD-SP20-2F	CTGTTGACATAGGCTTCGCA
	NYD-SP20-2R	TTTGCAGTTTACCAACATTCATT
3	NYD-SP20-3F	CATTAAAACCTGTCTCCAGAGA
	NYD-SP20-3R	TTTACACACTTGTCCATTGTAAGG
4a	NYD-SP20-4aF	TGGAATTTATTTTCAGGCGCT
	NYD-SP20-4aR	ATCTCGCAACAAAGAAACCG
4b	NYD-SP20-4bF	CTAGCCCTCTCATTGTGCAG
	NYD-SP20-4bR	TTAAGAAACCATATAGAAACAAATGC
5	NYD-SP20-5F	GGGAAGAATGAAGAGGAAGACA
	NYD-SP20-5R	CACCGGTATGTGGAGAAAT
6	NYD-SP20-6F	TTGCTTTCAAATTGAGAAATTAAGA
	NYD-SP20-6R	TCAAGTTTTTAAACCTCCAAA
7	NYD-SP20-7F	TGTGGATAATGGAAAGAAGAAA
	NYD-SP20-7R	TTTACAACCTAGATCTGCTTATTTGTCT
8	NYD-SP20-8F	TTGGACCTAAAAAGTAAAGAGGC
	NYD-SP20-8R	AATCACTGGGCTGAGGATTG
9	NYD-SP20-9F	GGGAGAGAAGAGGCTCCAGT
	NYD-SP20-9R	GATTCCCGCGCTTCTTC

3.5. Fine Chemicals

Deoxyribonucleoside triphosphates (dNTPs) were purchased from Promega (USA). PCR products to be sequenced were purified by using QIAquick PCR Purification Kit (Qiagen, Germany) according to the recommendations of the supplier.

3.6. Equipment

Autoclave	:	Model MAC-601, Eylea (Japan) Prior-Clave (UK)
Balance	:	Electronic Balance, GecAvery (UK) BJ210C Precisa (Switzerland)
Centrifuges	:	Centrifuge DuPont Instruments (USA) 5415C Eppendorf (Germany) Mini Spin Plus Eppendorf (Germany)
Deep Freezers	:	-20°C Bosch (Germany) -20°C AEG (Turkey) -70°C GFL (Germany)
Documentation System	:	BioDOC Video Documentation system Biometra (Germany) Gel Doc 2000 Bio-Rad (USA)
Electrophoretic Equipment	:	Horizon 58, Model 200 BRL (USA) Sequi-Gen Sequencing Cell Bio-Rad (USA) DCode Universal Mutation Detection System Bio-Rad (USA) Wide Mini Sub Cell Gt Bio-Rad (USA) Primo Minicell Thermo EC (USA)

Gel Drier	:	Biometra (USA)
Incubators	:	Plus Series Gallenkamp (Germany) Orbital Gallenkamp (Germany)
Magnetic Stirrer	:	Hotplate Magnetic Stirrer, HS31 Chiltern (UK) MR 3001 Heidolph (Germany)
Minishaker	:	IB InterMed (Denmark) Rotamax Heidolph (Germany)
Ovens	:	80°C Heraeus (Germany)
Power Supplies	:	Power Pac Model 3000 Bio-Rad (USA) Fotoforce 250 Fotodyne (USA)
Refrigerator	:	4°C Arçelik (Turkey)
Spectrophotometer	:	Lambda 3 UV/VIS Perkin-Elmer (USA)
Thermocyclers	:	MJ Research (USA) Techne Progene (UK) Thermal Reactor TR1 Hybaid (UK)
Transilluminators	:	Chromato-Vue Transilluminator, Model 1 TM-20UVP (USA) Fluorescent Table Consort (Belgium)
Vortex	:	Vortexmixer VM20 Chiltern Scientific (USA) REAX top Heidolph (Germany)

Waterbath	:	Water Bath D 3162 Köttermann Labortechnik (Germany)
Water Purification System	:	Millipore Elix 3 Millipore (France)

3.7. Electronic Database Information

The database information on physical maps of polymorphic markers was gathered from National Center for Biotechnology Information NCBI web site (<http://www.ncbi.nlm.nih.gov>) and utilized for fine mapping of known loci. The database used to assess candidate genes for the disorders under investigation was NCBI Map Viewer (<http://www.ncbi.nih.gov/mapview>). Information about genes was gathered from NCBI Entrez Gene (<http://www.ncbi.nih.gov/entrez/query.fcgi?db=gene>).

The database used for gathering information about the disorders was Online Mendelian Inheritance in Man (<http://www.ncbi.nlm.nih.gov/Omim>).

Primer design and alignment of sequenced exons were carried out with the PRIMER3 and CLUSTALW softwares, respectively, which were available at The Biology WorkBench (<http://workbench.sdsc.edu>).

The database used for gathering the software for linkage analyses and pedigree drawings was the Laboratory of Statistical Genetics at Rockefeller University (<http://linkage.rockefeller.edu>).

4. METHODS

4.1. DNA Isolation

Ten ml of a peripheral blood sample drawn into a sterile vacutainer tube containing K₂EDTA as anticoagulant was transferred to a 50 ml falcon tube, and thirty ml of cold cell lysis buffer was added and mixed thoroughly. The mixture was kept at 4°C for 15 minutes (min) to allow plasma membrane to be digested. It was centrifuged at 5000 revolution per minute (rpm) at 4°C for 10 min to collect leukocyte nuclei. The supernatant containing cell debris was discarded. The leukocyte nuclei were suspended in 10 ml cell lysis buffer. The centrifugation step was repeated. The supernatant was discarded, and the nuclei were resuspended in 5 of nuclei lysis buffer by vortexing until all clumps dissolved. Then, into this mixture, 80 µl of 10 per cent SDS and 50 µl of Proteinase K (20 mg/ml) were added and mixed gently by rolling and rocking the tube by hand. When the solution became jelly-like, the sample was incubated either at 56°C for 3 hours (hr) or at 37°C overnight to degrade the proteins. In order to get rid of the protein residues, 1.7 ml of 9.5 M NH₄Ac was added, and the sample was shaken vigorously to salt out proteins. Then, the sample was centrifuged at 5000 rpm at room temperature for 25 min. The supernatant was transferred into a clean 50 ml Falcon tube. Two volumes of ethanol were added to precipitate out the DNA. After DNA threads become visible and collected at the surface, the DNA was fished out carefully with a micropipette tip and transferred into a 1.5 ml Eppendorf tube. Residual ethanol was air-dried, and, depending in its estimated quantity, the DNA was dissolved in 200-1000 µl of TE buffer and stored at -20°C.

4.2. Linkage Analyses

Some of the Seckel families had already been mapped roughly in our laboratory. Those data were further confirmed, and the investigated loci were refined with additional markers. For analysis, commercial microsatellite markers were used whenever available; otherwise, new markers were designed from repeat sequences. The properties of both the designed and commercial markers (designation, fragment length, optimum annealing

temperatures and PCR buffer) are presented in Tables 4.1, 4.2, 4.3, and 4.4. The sequences of the amplified regions of designed markers are presented in Table 4.5.

To narrow down gene locus SCKL2, which is 15 Mb and at chromosome 18p11.31-q11.2, fine mapping was performed by using 22 microsatellite markers. In addition, all families were analyzed with 17 markers around SCKL3, which was identified by Kılınç [43]. Two new Seckel families that were not investigated by Kılınç [43] were tested for linkage to SCKL1. The patients were genotyped with marker D3S3694 that was closest to the *ATR* gene. In addition, *ATR* interacting and other selected DNA repair gene loci were analyzed for linkage with markers nearby. Regions where homozygosity was detected were analyzed further with new markers.

In azoospermia studies, the locus identified by Kılınç [43] was refined with additional markers using the large family. In addition, ten unrelated individuals were screened for linkage to this locus using the markers at the locus.

Polymerase chain reactions for microsatellite markers were carried out in a total volume of 11 μ l containing 1 X PCR buffer, 300 nM of each primer pair, 0.2 mM of dNTP, 125 ng of genomic DNA, 0.15 U Taq DNA polymerase, and sufficient dH₂O to complete the volume. After an initial denaturation step of 3 minutes at 95°C, 40 cycles of 30 sec denaturation at 94°C, 1 minute annealing at optimal temperature, and 30 sec elongation at 72°C, and a final elongation step at 72°C for 7 minutes were performed.

PCR products were mixed with 10 μ l of Stop solution and kept at -20°C until use. After incubating at 95°C for 5 minutes for denaturation, they were loaded onto an 8 per cent denaturing polyacrylamide gel as described in Sections 4.3.3. and 4.3.4. Alleles were resolved under constant power of 35 W for 1 to 3 hours and visualized by silver staining as described in Section 4.3.8.

Table 4.1. Properties of markers used for homozygosity mapping at gene locus SCKL2.

Designed markers are in italics

Marker Name	Product size reported (bp)	Repeat Type	Annealing Temperature ($^{\circ}\text{C}$)	Buffer Preference
D18S843	179-191	TRI	52	Boehringer
D18S464	283-291	DI	51.5	Boehringer
D18S1153	203-225	DI	59-54 touch-down	Boehringer
D18S1158	102-108	DI	53.5	Boehringer
D18S542	182-194	TETRA	53	Boehringer
D18S453	135-165	DI	53	Boehringer
D18S869	189	TETRA	53	Boehringer
D18S1107	114-128	DI	53	Boehringer
D18S975	153	TETRA	52	Boehringer
D18S866	202	TETRA	53	Boehringer
<i>D18SS03</i>	159	DI	51.6	Boehringer
D18S66	244-262	DI	54.5	Boehringer
D18S819	253	TETRA	51	Boehringer
D18S1151	211-219	DI	51	Boehringer
<i>D18SS04</i>	165	DI	53	Boehringer
<i>D18SL260</i>	225	TETRA	62-52 touch-down	Boehringer
D18S877	137	TETRA	54	Boehringer
<i>D18SL267</i>	173	TETRA	60-52 touch-down	Boehringer
D18S965	157	TRI	54	Boehringer
D18S847	212-236	TETRA	52	Boehringer
D18S535	150	TETRA	54	Boehringer
D18S858	191	TRI	54	Boehringer

Table 4.2. Properties of markers used for homozygosity mapping at gene locus SCKL3

Marker Name	Product size reported (bp)	Repeat Type	Annealing Temperature (°C)	Buffer Preference
D14S139	255	TETRA	52	Boehringer
D14S991	153-168	DI	52	Roche
D14S745	116	TETRA	52.5	Roche
D14S274	114-134	DI	54	Roche
D14S601	223	TRI	53	Boehringer
D14S994	214-226	DI	51	Boehringer
D14S586	214	TETRA	52	Boehringer
D14S1429	186-198	TETRA	54.5	Boehringer
MNATIntron	275	SNP	54	Boehringer
D14S592	237	TRI	52.5	Boehringer
D14S997	204-220	DI	54	Boehringer
D14S290	233-253	DI	51.5	Boehringer
D14S982	155-169	DI	52	Roche
D14S1069	200-220	DI	52	Roche
D14S119	169	TETRA	55	Boehringer
D14S588	117-141	TETRA	53	Boehringer
D14S77	203-249	DI	53	Boehringer

Table 4.3. Properties of markers used for homozygosity mapping at candidate DNA repair gene loci

Marker Name	Product size reported (bp)	Repeat Type	Annealing Temperature ($^{\circ}\text{C}$)	Buffer Preference
D3S2390	226-254	TETRA	60	Roche
D3S4545	237	TETRA	54	Roche
D3S2409	115-127	TRI	52.5	Roche
D3S3624	134-254	DI	51	Roche
D5S2500	155	TETRA	54	Roche
D5S629	233-253	DI	52	Roche
D5S1725	189	TETRA	53	Roche
D7S1823	230	TETRA	53	Roche
D8S1110	262-286	TETRA	54	Roche
D8S1460	301	TETRA	54	Roche
D8S1796	211-233	DI	60-54 touch-down	Roche
D9S910	107	TRI	60-55 touch-down	Roche
D11S1984	166-206	TETRA	52	Roche
D11S4464	237	TETRA	52.2	Roche
D11S968	150	DI	52	Roche
D11S2371	200	TETRA	55	Roche
D12S372	188	TETRA	55	Roche
D12S395	223-243	TETRA	54	Roche
D12S375	175-176	TETRA	52	Roche
D12S1025	186-197	TETRA	53	Roche
D12S1291	177-213	TETRA	53	Roche
D12S1601	193-203	DI	53	Roche
D12S1043	162	TRI	53	Roche
D12S303	295	TETRA	55	Roche
D12S828	286	TETRA	53	Roche
D12S376	148	TETRA	53	Roche
D12S56	122	DI	52	Roche
D14S1426	132-157	TETRA	52	Roche
D13S285	92-106	DI	52	Roche
D13S796	167	TETRA	52	Roche
ACTC	68-239	DI	62-55 touch-down	Roche
D17S1290	170-210	TETRA	53	Roche
D17S1293	262-290	TETRA	53	Roche
D22S683	168-214	TETRA	53	Roche
D22S689	202-226	TETRA	53	Roche

4.3. Candidate Gene Approach

To detect mutations in genes, exons and flanking sequences were analyzed by Single Strand Conformational Polymorphism (SSCP) analysis and direct sequencing. The basis of SSCP technique is that single strand DNA under non-denaturing conditions assumes a secondary structure based on its primary sequence. Thus, any single nucleotide change in its primary sequence is expected to result in a conformational change and alters the migration pattern in SSCP gels. The efficiency of the technique is estimated to be about 80 per cent.

4.3.1. Amplification of Candidate Genes by PCR

After investigating the databases, genes within the regions of interest were evaluated and candidate genes selected. Oligonucleotide primers were designed to amplify each exon and flanking sequences in the candidate genes *MNAT1*, *P2RX1*, *RAD51L1* and *NYD-SP20*. Amplified fragments were analyzed by SSCP and direct sequencing for any base change. In primer design, the general strategy is that product size must be shorter than 300 bp for optimal SSCP analysis but must include flanking intronic regions, which are important for RNA splicing. PCR reactions were carried out in a total volume of 25 μ l which contains 1 X PCR buffer, 0.2 mM of dNTP, 350 nM of each primer pair, 125 ng of genomic DNA, 0.2 U of Taq DNA polymerase, and sufficient dH₂O to complete the volume. Following an initial denaturation step of 4 minutes at 95°C, 40 cycles of 30 sec denaturation at 94°C, 1 minute annealing at optimal temperature, and 45 sec elongation at 72°C and a final elongation step at 72°C for 10 minutes were performed. Optimal PCR conditions for candidate genes are presented in Tables 4.6, 4.7, 4.8, and 4.9.

4.3.2. Analysis of Amplification Products on Agarose Gels

After amplification, PCR products were analyzed for extent of amplification. Five to seven μ l of PCR products were mixed with 1-1.5 μ l of 10 X loading buffer and loaded onto a 2 per cent agarose gel containing 3 μ g/ml of ethidium bromide, and run in 0.5 X TBE at 150-175 volts for 8-10 minutes. DNA bands were visualized under UV light transilluminator.

4.3.3. Preparation of Denaturing Polyacrylamide Gels

A 40 cm long polyacrylamide gel was cast. The apparatus was assembled using 0.4 mm spacers and combs. Thirty-five ml of 8 per cent Instagel was poured into a 50 ml beaker, mixed with 300 μ l of 10 per cent APS and 30 μ l of TEMED, and immediately poured between the glass plates of the apparatus. The sharks-tooth-comb was placed in an inverted position. The gel was allowed to polymerize for at least forty-five minutes before use.

4.3.4. Denaturing Polyacrylamide Gel Electrophoresis

As electrophoresis buffer, 1 X TBE pre-heated to 55-60°C in a microwave oven was used. The gel was first prerun in the hot buffer, prior to loading the samples, for 10-15 minutes in order to allow the gel to heat up uniformly to 45-50°C. During the prerun, samples were prepared for loading by first mixing with 10 X Stop buffer in a 1:1 ratio, then denaturing at 95°C for 4-5 minutes in a heat block and immediately placing on ice. Next, the comb was removed, and the upper part of the gel was cleaned from urea and gel particles by squirting electrophoresis buffer. The comb was placed in the correct orientation to form loading wells. Depending on their band intensities in agarose gels, 3 to 6 μ l of each sample was loaded into the slots. The gel was run at a constant power of 35 W for 1-3 hours to resolve the alleles. In order to visualize the fragments, the gel was silver stained as explained in Section 4.3.8.

4.3.5. Single Strand Conformational Polymorphism Analysis

Within the scope of this study, *MNAT1*, *P2RX1*, and *RAD51L1* were first analyzed by SSCP. Each of these candidate genes was analyzed by using PCR products of an affected individual, a heterozygote carrier, and an independent control individual who was expected not to carry any mutation. In order to increase the sensitivity of the mutation detection method, analysis was performed at various conditions. The different conditions were non-denaturing gel matrices with or without 5 per cent glycerol at 4°C or 20°C of electrophoresis temperatures. When a variant SSCP pattern was observed, all family members were analyzed with SSCP in order to assign the type of base change as a

mutation or polymorphism. After SSCP analysis, the fragments were subjected to direct sequencing analysis also in order to detect any base change that might have been missed by SSCP analysis.

4.3.6. SSCP Gel Preparation

The gel glasses were 20 cm x 20 cm in size, and the apparatus was assembled using 0.75 mm spacers. Six ml of 40 per cent acrylamide stock was mixed with 1.8 ml of 10 X TBE, and the volume was completed to 30 ml by adding sufficient dH₂O. Into this solution, 300 µl of 10 per cent APS and 30 µl of TEMED were added and mixed thoroughly, and the solution was poured between the glass plates using a plastic beaker. A 20-well comb was inserted immediately after the gel was poured, and the gel was left to polymerize for at least an hour. After polymerization, the gel was cooled at 4°C for at least fifteen minutes before use.

4.3.7. SSCP Electrophoresis

The samples were mixed in 1:1 ratio with 10 X Stop buffer, denatured at 95°C in a heat block for 4-5 minutes, and immediately placed on ice. 7 to 10 µl of each sample was loaded. Electrophoresis was carried out in 0.6 X TBE buffer at 20°C at a constant power of 6 W for 12-18 hours in the BioRad DCode™ Universal Mutation Detection System. Finally, in order to visualize the bands, the gel was silver stained as explained in Section 4.3.8.

4.3.8. Visualization

After electrophoresis was complete, the glass sandwich containing the gel was removed from the apparatus and let cool down for a few minutes. The glass plates were separated, allowing the gel to remain on one of them. In PAGE, the gel was transferred into Buffer B by the help of a filter paper. A filter paper was cut to the size of the gel area that was assessed to contain the fragments. The paper was laid onto the gel and used to lift and detach the gel from the glass plate, whereas in SSCP, a filter paper was not used, but the gel was placed in Buffer B with the glass plate. In PAGE, the gel was kept with the

filter paper with great care because the gel was very thin and prone to tearing. The gel was shaken in Buffer B for 10-12 minutes. Buffer B was discarded, and the gel was washed twice briefly in dH₂O to remove excess silver. As a next step, Buffer C was poured on the gel, and the gel incubated until bands appeared. Finally, Buffer C was discarded, Buffer D was added, and the gel soaked in it for 10 minutes to stop the reaction. The gel was transferred into a transparent folder for storage and documentation with the aid of a paper towel [45].

4.3.9. DNA Sequence Analysis

In the scope of this study, *MNAT1*, *P2RX1*, *RAD51L1*, and *NYD-SP20* were subjected to sequence analysis. For preparation of samples for sequencing, PCR reactions were carried out in 100 µl total volume, and the products purified from the dNTPs, excess primers, and genomic DNA by using a QIAquick-spin column. In this protocol, firstly, 500 µl of buffer PB was added to the PCR product and mixed well, and the mixture was applied to a QIAquick-spin column and centrifuged at 13 rpm for 1 minute. During this step, amplification products bind to the filter inside the QIAquick-spin column. The other molecules were washed away by adding 750 µl buffer PE and centrifuging at 13 rpm for 1 minute. PE buffer was discarded from the lower collector tube and centrifugation repeated. The column was placed into a sterile 1.5 ml microcentrifuge tube, and forty µl of dH₂O was applied to the column and left for 1 min. After the final centrifugation step of 1 min at 13 rpm, purified DNA fragments with sizes between 100-10000 bp were eluted. The purity of DNA fragments was confirmed using polyacrylamide gel electrophoresis.

Pure DNA fragments were later either sent to Harvard University DNA Sequencing Facility or given to Dr. Durişehvar Ünal for sequence analysis. DNA sequencing instrument used was Amersham Biosciences ABI 310 gene analyzer.

Table 4.6. PCR conditions to amplify *MNAT1*

Exon	Size (bp)	Annealing Temperature (°C)	Buffer Preference
MNAT1	309	58.2	Boehringer
MNAT2	331	58.4	Boehringer
MNAT3	266	58.5	Boehringer
MNAT4	288	58.7	Boehringer
MNAT5	275	58.2	Boehringer
MNAT6	348	57.0	Boehringer
MNAT7	297	58.0	Boehringer
MNAT8	355	59.0	Boehringer

Table 4.7. PCR conditions to amplify *RAD51L1*

Exon	Size (bp)	Annealing Temperature (°C)	Buffer Choice
RAD51L1-1	336	56	Roche
RAD51L1-2	282	56	Roche
RAD51L1-3	291	56	Roche
RAD51L1-4	289	59	Roche
RAD51L1-5	240	56	Roche
RAD51L1-6	310	57	Roche
RAD51L1-7	224	57	Roche
RAD51L1-8	274	57	Roche
RAD51L1-9	226	57	Roche
RAD51L1-10	255	56	Roche
RAD51L1-11	199	56	Roche

Table 4.8. PCR conditions to amplify *P2RX1*

Exon	Size (bp)	Annealing Temperature (°C)	Buffer Preference
P2RX1	318	50	Boehringer
P2RX2	323	56	Boehringer
P2RX5	331	56	Boehringer
P2RX4	343	56	Boehringer
P2RX5	310	50	Boehringer
P2RX6	346	50	Boehringer
P2RX7	348	61-50 touch-down	Boehringer
P2RX8	322	54	Boehringer
P2RX9	281	56	Boehringer
P2RX10	273	56	Boehringer
P2RX11	283	54	Boehringer
P2RX12	289	61-50 touch-down	Boehringer

Table 4.9. PCR conditions to amplify *NYD-SP20*

Exon	Size (bp)	Annealing Temperature ($^{\circ}$ C)	Buffer Choice
NYD-SP20-1a	278	57	Roche
NYD-SP20-1b	350	65-61 touch-down	Roche
NYD-SP20-2	190	57	Roche
NYD-SP20-3	299	57	Roche
NYD-SP20-4a	276	57	Roche
NYD-SP20-4b	292	65-61 touch-down	Roche
NYD-SP20-5	300	57	Roche
NYD-SP20-6	251	57	Roche
NYD-SP20-7	308	65-59 touch-down	Roche
NYD-SP20-8	315	57	Roche
NYD-SP20-9	231	59	Roche

5. RESULTS

The principle in homozygosity mapping is to uncover a homozygous genomic block shared by affected individuals. Such a block supposedly originates from a single common ancestor and possibly harbors the gene responsible for the disorder. In the framework of linkage studies, two autosomal recessive disorders were investigated by homozygosity mapping. Additionally, computerized lod score analysis was performed where needed using the data obtained from the linkage studies. The genes within the identified homozygous blocks were explored by using Mapviewer tool of NCBI database (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/Mapviewer>) and then the functions of genes were analyzed in detail in order to select the most plausible candidate genes that would be implicated in the etiology of the diseases under investigation.

Later, candidate genes within homozygous regions were selected and investigated for possible mutations.

5.1. Linkage Analysis

5.1.1. Seckel Syndrome

Seckel is a rare autosomal recessive disorder, but as consanguineous marriages are common in Anatolia and those families have increased incidence of such disorders, fifteen families could be collected, which were of great importance in this study. Consanguinity was reported in eight: families 1, 2, 3, 5, 8, 10, 11 and 12. In the remaining, it was highly probable that parents were distantly related, because they originated from the same small Anatolian villages or towns. Five of the families had two affected sibs each (families 1, 2, 3, 4, and 8), while the remaining had one each. A total of 20 patients and 51 family members were studied. Table 1.2 presents the families studied.

Prior to this study, initial analyses had been performed in 13 families in our laboratory. The data are summarized in Section 1.3. In the present study, all families were

investigated in detail at known loci, and additional markers were employed to refine and narrow down three of those loci.

Fine mapping was performed for 14 families using 22 markers within and around SCKL2 locus and with 17 markers around SCKL3 locus. For the first 14 families, 4 ATR related and 15 other selected candidate DNA repair gene loci were tested for linkage using nearby microsatellite markers. For markers for which homozygosity was observed, parents were analyzed in order to assess informativeness of the marker for the family and unaffected sibs to make sure that they do not share the homozygous genotype.

Firstly, linkage studies were performed around SCKL2 locus between markers D18S843 and D18S866, which were 14,162 Mb apart. For 10 of 15 families, linkage was excluded, as the families were not indicative of linkage to this locus. The linkage results are given in Table 5.1. In family 15, homozygosity was observed in a maximal 3,635 Mb region between markers D18S866 telomerically and D18S-S03 centromerically within the SCKL2 locus and in a minimal critical region of 1,824 Mb between D18S1107 and D18S866. The linkage results of this family are presented in Figure 5.1. The asterisk indicates a crossover in S15. In addition, this family was found homozygous for three consecutive markers D18S819, D18S1151, and D18S04 in a maximal region spanning 1,982 Mb between markers D18S66 and D18SL260 and in a minimal critical region of 1,243 Mb. For families 9 and 12, homozygosity was detected for markers D18S464, D18S1153, D18S1158, and D18S542 in a minimal critical region of 1,531 Mb. This homozygous block was flanked by markers D18S843 and D18S453 which were 4,301 Mb apart. As the two families shared homozygosity in this region, it is possible that they have a defect in a common gene in this region. The families were too small for lod score analysis, thus computational analysis was not performed. Also, family 11 was found homozygous for markers D18SL260, D18S867, D18SL267 and D18S965, which were at flanking centromeric region of the identified SCKL2 locus and in a minimal critical region of 0,98 Mb wide. This block was flanked by markers D18S-S04 and D18S847 which were 2,253 Mb apart. Examples of denaturing polyacrylamide gel results for markers used in linkage analyses at SCKL2 are presented in Figures 5.2, 5.3, and 5.4. The numerals on gel pictures in figures point out to various alleles.

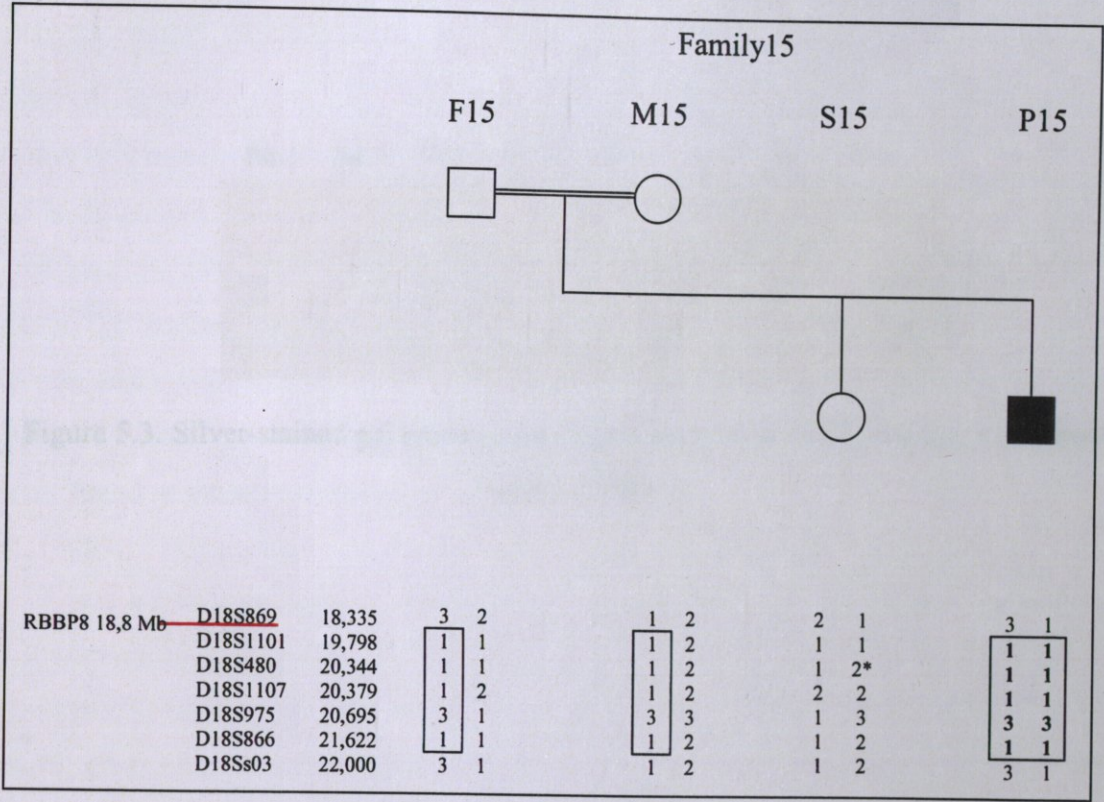


Figure 5.1. Linkage results of family 15 at SCKL2

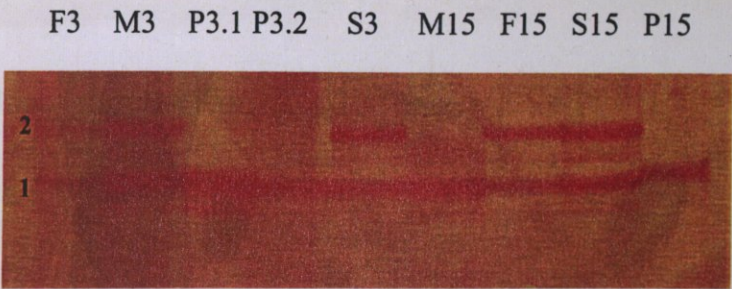


Figure 5.2. Silver-stained gel showing the genotypes of some Seckel family members for marker D18S866

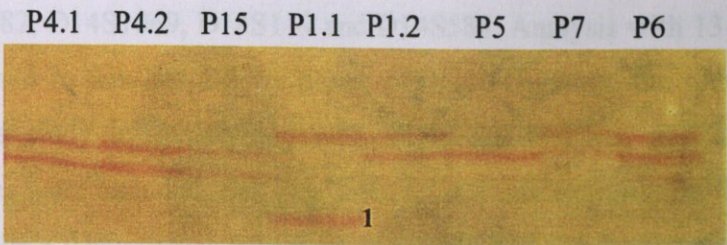


Figure 5.3. Silver-stained gel showing the genotypes of some Seckel family members for marker D18S877

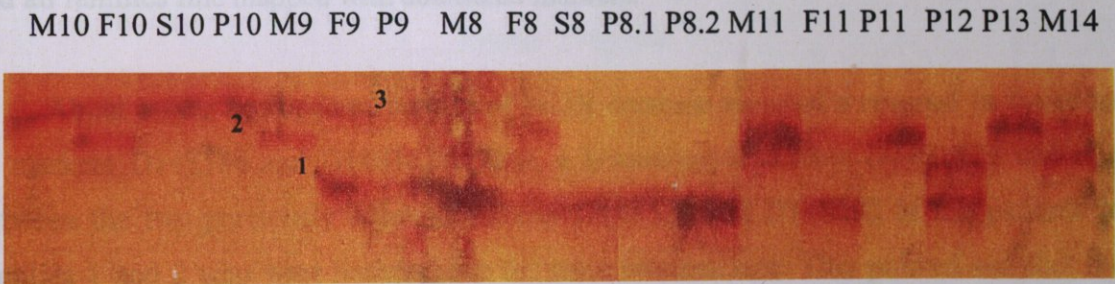


Figure 5.4. Silver-stained gel showing the genotypes of some Seckel family members for marker D18S965

Linkage analysis for all seventeen families was performed at SCKL3 using 17 markers. Initially in family 3, the defined locus was analyzed with additional markers D14S745, D14S274, MNATIntron5 polymorphism, and D14S1069. A homozygous block of 19,721 Mb was identified. The homozygous region included markers D14S745, D14S274, D14S601, D14S994, D14S586, D14S1429, MNATIntron5, D14S592, D14S997, D14S290, D14S982, D14S1069, D14S119 and D14S588. Analysis with 13 of the markers had been performed in families 1-3 by Kılınç [43]; nevertheless, the previous data was confirmed in this study by genotyping with all of the markers, and additionally the remaining families were included. Samples from all patients were run on the same PAGE gels so that the allele designations would be consistent and the frequencies of the alleles could be determined. This aided to identify rare alleles, which were highly informative when found in the homozygous state.

In the previous studies, 7 of the families were roughly mapped around SCKL3 [43]. The patients in families 1-3 were found to share homozygosity for markers D14S586 and D14S592, in family 5 for only D14S586, and in families 4 and 7 for only D14S592. In this study, other families were tested for linkage to the locus, the previous results confirmed, and all families fine mapped with additional markers.

About a year ago updated version of NCBI database reported a reversal in the order of markers D14S586 and D14S1429. Thus, a heterozygous marker D14S1429 came in between the two previously adjacent markers D14S586 and D14S592 in the patients in families 1 and 2 who were homozygous for these two markers. This pointed out that the real SCKL3 locus might not be centered around marker D14S592 but be somewhere else in the large homozygous region in family 3. Therefore, all families were screened for any significant linkage within this large homozygous region. The results are presented in Table 5.2. In families 6 and 12, homozygosity for markers D14S139 and D14S745 was identified. The region was 3,954 Mb and between D14S989 telomerically and D14S274 centomerically. Family 6 was found not linked to SCKL1 or SCKL2, and as its size was small, a genome-wide scan had not been performed to identify the gene locus. The patients exhibited homozygosity also for 3 consecutive markers D14S982, D14S1069, and D14S119 at SCKL3. Homozygosity was in a minimal critical region of 2,702 Mb and a maximal 6,688 Mb region between D14S290 and D14S588. In family 12, a different small

homozygous region was identified between D18S843 and D18S453 (Table 5.1). A genome-wide scan had not been performed as this family was also small. In family 7, a homozygous region of 7,659 Mb was identified, delineated by markers MNATIntron5 polymorphism and D14S119 in a minimal critical region of 6,992 Mb. The patient was homozygous for 5 consecutive markers: (tel) – D14S592, D14S997, D14S290, D14S982, and D14S1069 – (cen). Homozygosity in patient 6 and patient 7 included D14S1069 that was within *RAD51L1*. The remaining 10 families were excluded to link to SCKL3, assuming that the current information in NCBI database was correct. Examples of genotyping with markers at SCKL3 are presented in Figures 5.5, 5.6, 5.7, and 5.8. The letter C denotes control individuals.

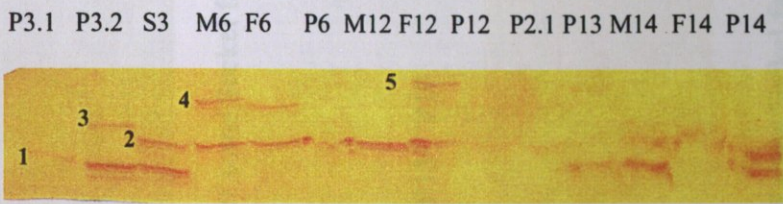


Figure 5.5. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S139

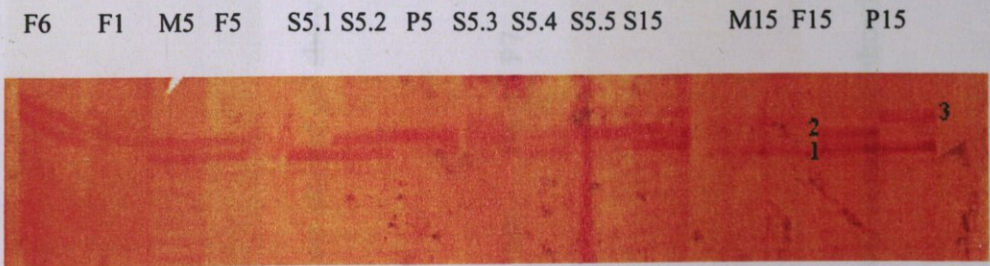


Figure 5.6. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S586

F10 M10 S10 P10 P5 P12 P13 P14 P10 P7 M3 F3 P3.1 P3.2 S3 P6 P2.1 P4.1 P9 P8.2 P1.1



Figure 5.7. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S119

M6 P6 P7 P15 P14 P13 P10 P9 C1 C2

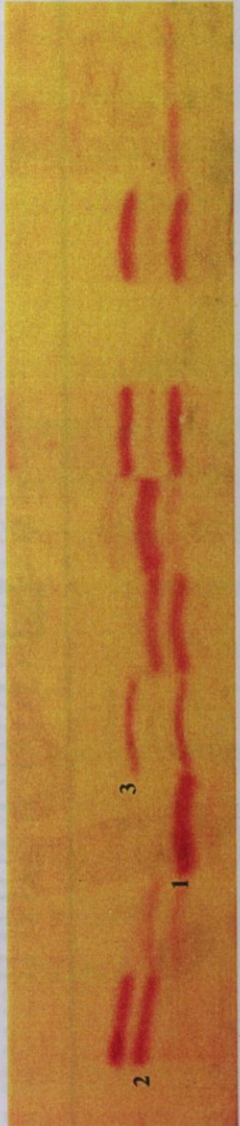


Figure 5.8. Silver-stained gel showing the genotypes of some Seckel family members for marker D14S1429

In addition to the three SCKL loci, 19 candidate genes were evaluated for linkage using a total of 35 microsatellite markers which were 0.2 to 10 Mb away from the genes, except for 3 genes that were analyzed with markers 10-15 Mb away. A list of the genes, their loci and the markers employed are presented in Table 1.1. Four of the genes coded for important proteins that were in the same pathway as ATR. The remaining 15 were genes implicated to play crucial roles in DNA repair. Examples of denaturing polyacrylamide gel results for markers used in linkage analyses are presented in Figures 5.9, 5.10, 5.11, 5.12 and 5.13. Families 1 and 3 were not included in the analysis because they already linked to SCKL1 and SCKL3, respectively. As part of linkage analysis to various genes, parents of homozygous patients were subsequently genotyped, unless affected siblings had different genotypes.

One of the genes investigated was *DNAPK*. Microsatellite markers D8S1460, D8S1716 and D8S1110 were used. In family 11, a minimum 6.3 Mb homozygous region was identified in the patient, and the parents were informative. Linkage results of Seckel families around this gene are presented in Table 5.3. Twelve of the families were excluded from linkage to the *LIG4* locus, and in the remaining family 11, parents were found to be noninformative for D17S1396 (Table 5.4).

Only the patient in family 7 was homozygous for the two markers close to *RAD17* and *XRCC4*. However, only one parent was informative (Table 5.5).

All families analyzed were excluded from linkage to *RAD51L3*, *LIG3*, and *RAD52B* using D17S1293, as no patient was found homozygous (Table 5.6).

Patients 12 and 15 were found homozygous around *RAD9* gene for the marker D11S2371 (Table 5.7).

The patients in families 9 and 11 were homozygous for D12S372 that was nearby the *RAD52* gene. But parents in family 11 were noninformative (Table 5.8).

Eleven of the families were found not to link to *XRCC2* using marker D7S1823. Patients in families 8 and 11 were found homozygous and parents informative for this marker (Table 5.9).

All families were excluded from RAD18 locus, as they were all found in the heterozygous state for marker D3S4545 (Table 5.10).

Linkage analysis to *RAD51C* gene using D17S1290 resulted in the exclusion of 12 of the families. Only the patient in family 7 was found homozygous for the marker, and the parents were informative (Table 5.11).

Patients in families 9 and 13 were homozygous for D14S1426 in the vicinity of *XRCC3*. Parents were informative in family 13, and only one parent was informative for family 9. Other patients were not found homozygous (Table 5.12).

Patients in families 6, 7, 11 and 13 were homozygous for marker D22S689, which was nearby *CHK2* gene. Families 6 and 11 were noninformative, while families 7 and 13 informative for only one parent. No patient with both parents informative was homozygous for the marker (Table 5.13).

After a crude analysis around *MDM2* for all families, linkage was excluded in 10 of the families. Families 8, 9 and 15 were indicative of linkage, as patients were homozygous for D12S375. Both parents were informative in families 9 and 15, and one parent in family 8 was informative. However, patients were not homozygous for the flanking markers (Table 5.14).

Analysis with D11S4464, which was in the vicinity of *CHK1* gene, revealed homozygosity in patients in two large families (2 and 5). However, when these families were tested with flanking markers (D11S968 telomericly and D11S968 centromericly), the patients were found heterozygous (Table 5.15).

All 13 families were excluded from linkage to RAD51 locus, as all patients were heterozygous for GATA50C03.

In sum, of the 13 families studied for linkage to candidate genes, family 7 was indicative of linkage to *RAD51C*, family 9 and 15 to *MDM2*, family 9 to *RAD52*, family 10 to *RAD17*, family 11 to *DNAPK* and *XRCC2*, and family 13 for *XRCC3*. One or both parents were noninformative in the following families for the markers close to the indicated genes: families 4 and 7 for *XRCC4*, family 8 for *RAD17*, family 9 for *XRCC3*, family 10 and 11 for *RAD52*, family 11 for *LIG4* and families 6, 7, 11, and 13 to *CHK2*. In addition, families 2, 5, 12 and 14 were not indicative of linkage to any of the gene loci tested. Thus, linkage to those loci could not be evaluated.

Table 5.14. Linkage results around *CDK2*, *CDK4* and *MDM2* for Seckel families whose patients were homozygous for D12S375

LOCUS	POSITION (Mb)	F8	M8	S8	P8.1	P8.2	F9	M9	P9	F15	M15	S15	P15
<i>CDK2</i>	54,6												
<i>CDK4</i>	56,4												
D12S1601	65,8	--	--	--	--	--	13	23	12	12	12	22	12
D12S1291	66,2	13	14	14	14	14	22	21	22	25	56	55	56
D12S375	67,2	55	56	56	55	55	51	52	55	1-	13	13	11
<i>MDM2</i>	67,5												
D12S1043	69,2	--	--	--	12	12	--	--	--	12	11	12	12
D12S1025	69,5	11	11	11	11	11	11	22	12	12	13	13	11
D12S828	73,4	22	12	22	12	12	--	--	--	22	12	22	12
D12S376	73,5	11	12	11	12	12	--	--	--	11	11	11	11

Table 5.15. Linkage results around *CHK1* for Seckel families whose patients were homozygous for D11S4464

LOCUS	POSITION (Mb)	M2	F2	P2.1	P2.2	S2.1	S2.2	M5	F5	P5	S5.1	S5.2	S5.3	S5.4	S5.5
D11S4464	123,1	35	36	33	33	35	36	26	23	22	23	12	23	36	23
D11S4158	124,4	--	--	23	23	--	--	--	--	13	--	--	--	--	--
<i>CHK1</i>	125														
D11S968	125,6	--	--	12	12	--	--	--	--	23	--	--	--	--	--

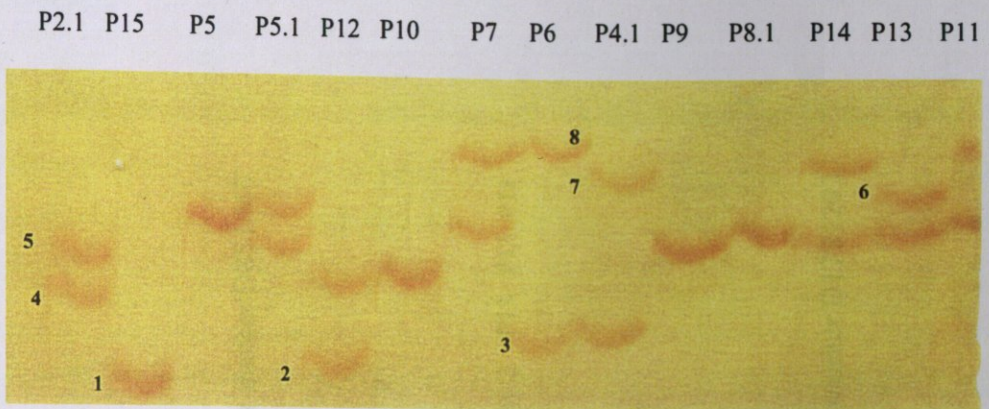


Figure 5.9. Silver-stained gel showing the genotypes of some Seckel family members for marker D12S375

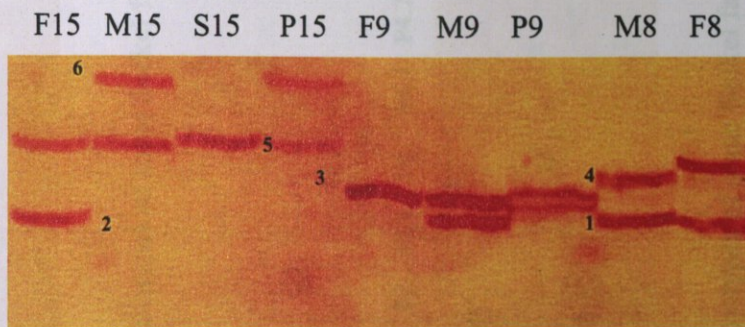


Figure 5.10. Silver-stained gel showing the genotypes of some Seckel family members for marker D12S1291

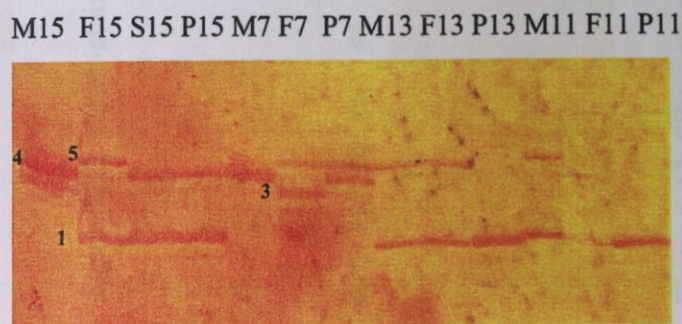


Figure 5.11. Silver-stained gel showing the genotypes of some Seckel family members for marker D8S1460

5.1.2. Azoospermia

A family with three boys with azoospermia and one unaffected was investigated in this study. Azoospermia was not associated with congenital bilateral absence of vas deferens, as opposed to that in CF patients. The initial analysis was performed by Kılınç [43], and CFTR gene was also screened for mutations. Two of the patients but not the third one was found to carry the novel mutation D891G. Also, no AZF microdeletions was found in the affected individuals. These data suggested that azoospermia in the family was associated with CF, and possibly another gene was involved. A genome-wide scan was performed, and a shared homozygosity in the patients was identified at chromosome 17 [43]. The gene locus was at chromosome 17p13 between markers D17S1866 and D17S973 in a maximal region of 8,92 Mb. Intriguingly, a heterozygous marker interrupted the maximal region.

In this study, an attempt was made to refine the locus with additional markers D17S525, D17S1152 and D17S796. In addition, 10 unrelated azoospermia patients were analyzed for linkage to this locus. In the interrupting region, affected individuals were found heterozygous also for D17S525. The physical locations of markers and genotypes are presented in Table 5.16. Linkage analysis for other 10 unrelated azoospermic individuals was not completed. In the initial analysis only one patient (patient 2) was found homozygous, and homozygosity included four consecutive markers (D17S1308, D17S926, D17S695, and D17S1529) in a minimal critical region of 0,427 Mb.

In this study, two-point linkage analysis was performed using the program MLINK and multi-point lod scores were computed using GENEHUNTER. The maximum two-point lod score obtained at the locus was 2.18 for markers D17S695 and D17S1533 at a recombination fraction of zero. Two point lod scores for each marker versus the disease locus are presented in Table 5.17. Seventeen-point lod score analysis resulted in a maximum lod score of 3.259 between markers D17S1308 and D17S695 at the location 0.3 Mb - 0.7 Mb on the map given in Figure 5.14. The lod score was higher than 3 also between markers D17S559 and D17S796.

Table 5.16 Haplotypes of azoospermia family and other azoospermia patients

Family		F1	M1	P1.1	P1.2	P1.3	S.1.1	S.1.2	P2	P3.1	P3.2	P4.1	P4.2	P5	P6	P7	P8	P9
Sex		m	f	m	m	m	m	f	m	m	m	m	m	m	m	m	m	m
LOCUS	POSITION (Mb)																	
D17S1866	0,125	22	21	22	22	22	22	21	13	11	11	11	11					
D17S1308	0,607	23	22	22	22	22	32	32	33	23	23	23	23	57	45	66	45	23
D17S926	0,615	22	21	22	22	22	22	21	22	22	22	12	12	12	11	11	13	13
D17S695	0,73	23	21	22	22	22	32	31	22	14	14	36	36	48	67	26	23	46
D17S1529	1,03	21	21	22	22	22	12	11	33	23	23	13	13			44	44	15
D17S1533	1,74	31	32	33	33	33	13	12	12									
D17S525	2,15	12	11	11	11	11	21	21	13					24	34	23		
D17S1152	3,509	22	12	21	21	21	21	22										
NYD-SP20	3,554-3,581																	
D17S919	3,89	21	31	23	23	23	13	11	34	22	22	13	13					
P2RX1	4,006-4,026																	
D17S1828	4,01	21	22	22	22	22	12	12	23	23	23	22	22					
D17S559	4,94	22	21	22	22	22	22	21										
D17S1854	5,61	22	21	22	22	22	22	22										
D17S513	5,88	11	12	11	11	11	11	11										
D18S796	6,45	11	11	11	11	11	11	11										
D17S973	8,92	11	11	11	11	11	11	11										
D17S974	10,46	32	42	34	32	34	24	24	44	34	33	23	23					
D17S1303	10,8	43	23	42	43	42	32	32										

Table 5.17. Two-point lod scores between azoospermia
and 16 markers at chromosome 17p13

Locus	Lod score at $\theta =$						
	0	.01	.05	.1	.2	.3	.4
D17S1866	0.97	0.95	0.86	0.75	0.51	0.29	0.1
D17S1308	1.27	1.24	1.13	0.99	0.69	0.39	0.14
D17S926	0.97	0.95	0.86	0.75	0.51	0.29	0.10
D17S695	2.18	2.13	1.92	1.66	1.14	0.64	0.22
D17S1529	2.14	2.09	1.88	1.62	1.09	0.59	0.18
D17S1533	2.18	2.13	1.92	1.66	1.14	0.64	0.22
D17S525	1.27	1.24	1.13	0.99	0.69	0.39	0.14
D17S1152	-4.02	-0.90	-0.26	-0.05	0.06	0.04	0.01
D17S919	-2.69	0.17	0.69	0.76	0.60	0.32	0.06
D17S1828	1.27	1.24	1.13	0.99	0.69	0.39	0.14
D17S559	0.88	0.89	0.87	0.97	0.56	0.31	0.11
D17S1854	0.60	0.59	0.54	0.49	0.35	0.19	0.07
D17S513	0.21	0.20	0.18	0.15	0.10	0.06	0.02
D18S796	0.21	0.20	0.18	0.15	0.10	0.06	0.02
D17S973	$-\infty$	-1.83	-0.60	-0.17	0.09	0.09	0.02
D17S974	$-\infty$	-1.86	-0.59	-0.18	0.05	0.06	0.00

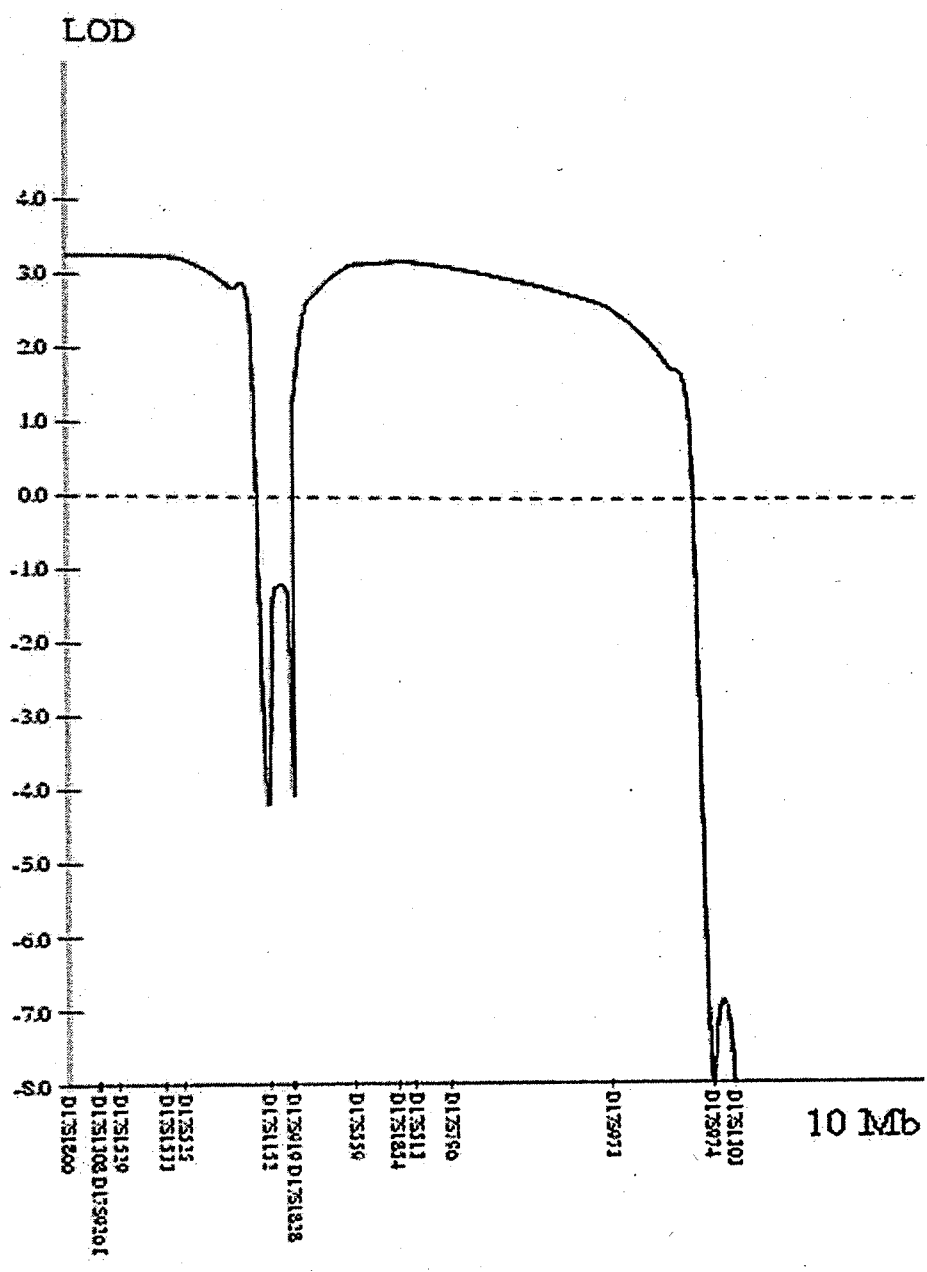


Figure 5.14 Results of seventeen-point lod scores at chromosome 17p13 for azoospermia using GENEHUNTER

5.2. Gene Analysis

5.2.1. *RAD51L1*

Two candidate genes at SCKL3 and two at the azoospermia locus were analyzed for mutations in relevant families. *RAD51L1* was noticed in the current databases. SSCP analysis of this gene were performed in all patients in families 3 and 6 and additionally 3

5.2.1. *MNAT1*

of these families. Later, sequencing of this gene was carried out for patient 3.1 and patient 6. No variation was detected.

In the initial results of Kılınc [43], *MNAT1* had been selected as a candidate gene responsible for SCKL3, as the gene contained the marker D14S592, where the maximum LOD score appeared [43]. In the present study, all eight exons of *MNAT1* were screened for mutations by SSCP in four Seckel patients from families 1, 2, 3, and 7, which were homozygous for marker D14S592. No mutation but a polymorphism (77205A/G) at intron 5 was detected, first by SSCP and then by direct sequencing. As two of the patients were found to have a heterozygous pattern for the fragment containing exon 5 and flanking regions, their sequence analysis was discontinued. Allele G was not reported in the databases. Out of fifteen controls analyzed for the variant, three were found in the homozygous and another three in the heterozygous state, resulting in 30 per cent frequency. The observation of the base substitution in several of the controls and parents both in heterozygous and homozygous state lead to the classification of this substitution as a polymorphism. An example of denaturing polyacrylamide gel results for this variant are presented in Figure 5.15. The letter C denotes control individuals.

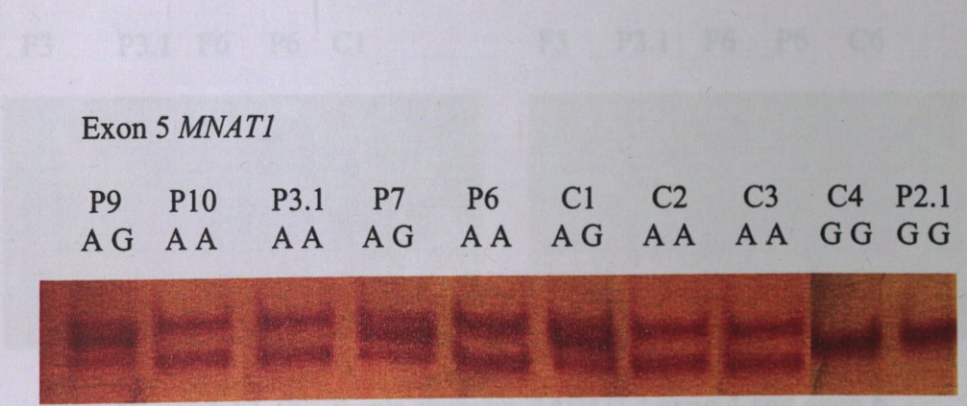


Figure 5.15. SSCP gel showing 77205A/G in *MNAT1*

5.1.2. *RAD51L1*

A better candidate gene *RAD51L1* was noticed in the current databases. SSCP analysis of this gene were performed in all patients in families 3 and 6 and additionally a parent from one of these families. Later, sequencing of this gene was carried out for patient 3.1 and patient 6. No variation was detected.

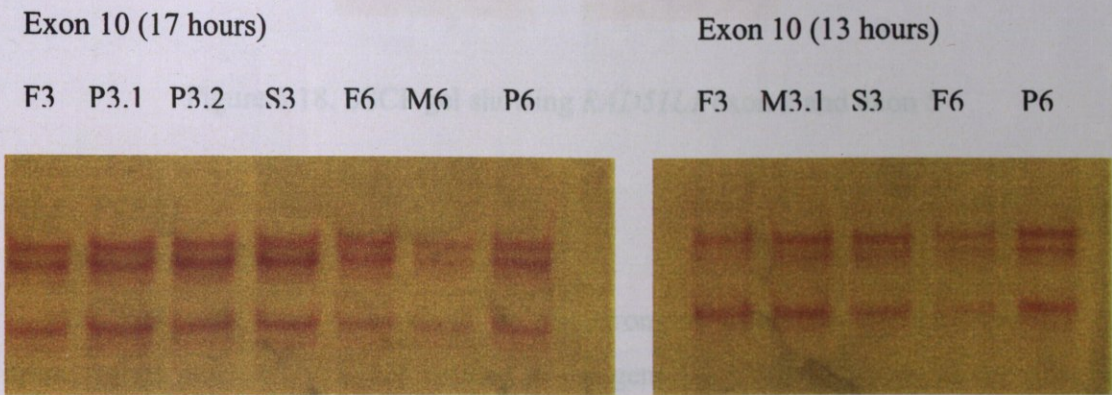


Figure 5.16. SSCP gels showing *RAD51L1* exon 10

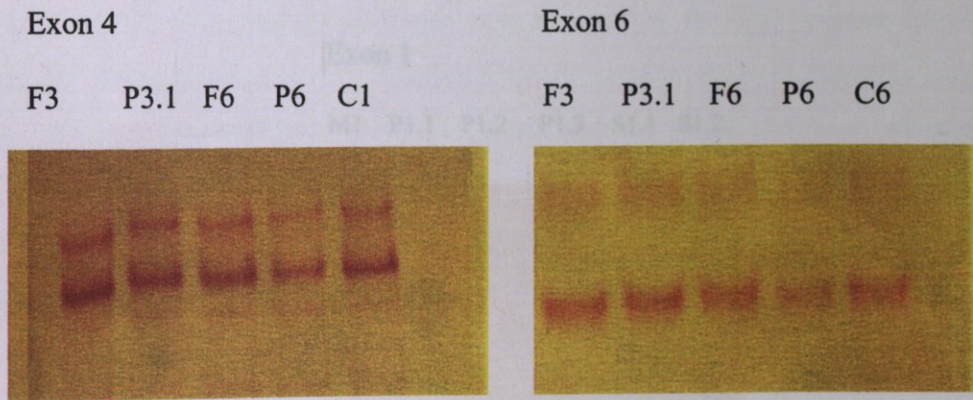


Figure 5.17. SSCP gels showing *RAD51L1* exon 4 and exon 6

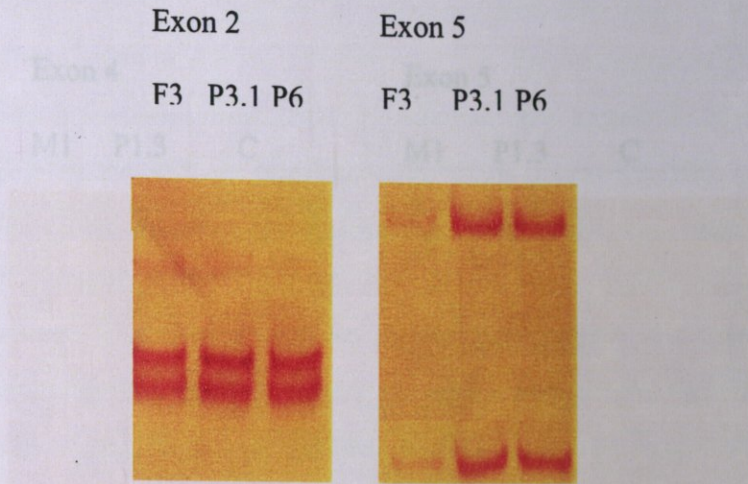


Figure 5.18. SSCP gel showing *RAD51L1* exon 2 and exon 5

5.1.3. *P2RX1*

P2RX1 gene was assessed to be the strongest candidate for azoospermia. The homozygous marker D17S1828 resided in the gene [43]. All 12 exons of the gene was analyzed by SSCP in all family members. In addition, sequencing was performed with individuals M1 and P1.3. But no variant was detected.

NYD-SP26 was another candidate gene that was at the azoospermia locus, but not within the homozygous region. It resided in the region between the two homozygous blocks. As databases still had no information about this gene, it was a very good candidate, the patients were sequenced. The samples from individuals M1 and P1.3 were used for direct sequencing of the gene. No variant was detected in this gene either.

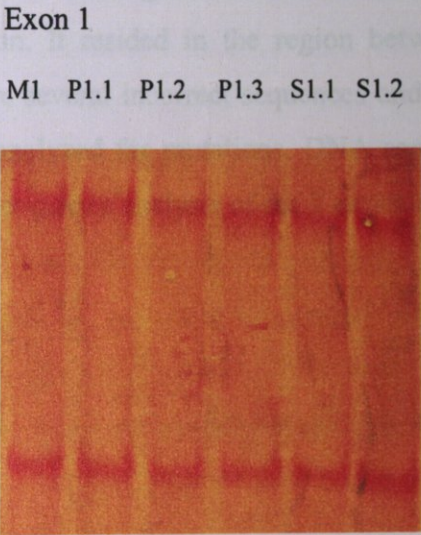


Figure 5.19. SSCP gel showing *P2RX1* exon 1

6. DISCUSSION

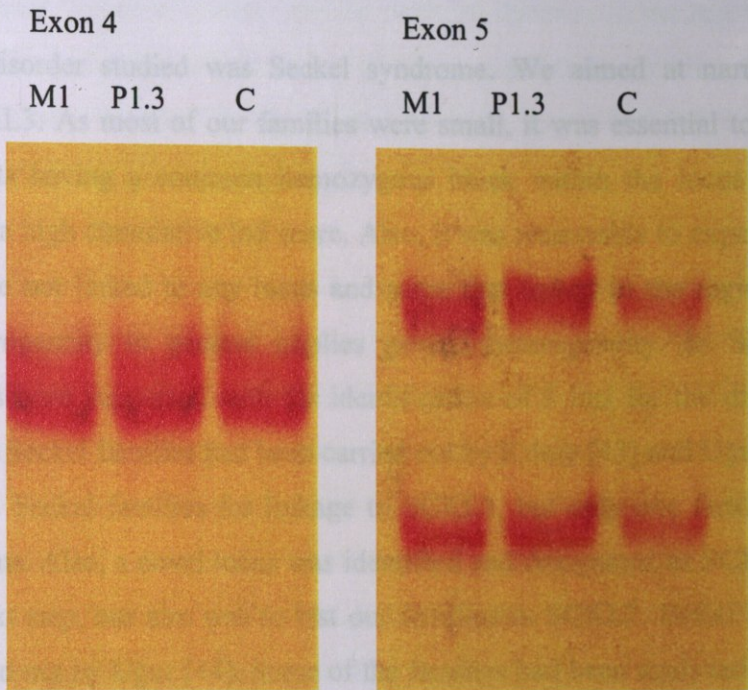


Figure 5.20. SSCP gel showing *P2RX1* exon 4 and 5

5.1.4. NYD-SP20

NYD-SP20 was another candidate gene that was at the azoospermia locus, but not within the homozygous region. It resided in the region between the two homozygous blocks. As databases still have several incorrect sequences and the gene was a very good candidate, the patients were analyzed for mutations. DNA samples from individuals M1 and P1.3 were used for direct sequence analysis of the 9 exons of the gene. No variant was detected in this gene either.

6. DISCUSSION

The first disorder studied was Seckel syndrome. We aimed at narrowing down SCKL2 and SCKL3. As most of our families were small, it was essential to find at least three-four patients having a common homozygous block within the locus to narrow it down and obtain a high cumulative lod score. Also, it was reasonable to expect to hit upon families that were not linked to any locus and some that linked to the regions analyzed. Phenotypic heterogeneity in general implies genetic heterogeneity. In Seckel genetic heterogeneity has been supported with the identification of 3 loci for the disorder. Initial mapping in some Seckel families had been carried out by Kılınç [43] and Uğur [44]. Kılınç [43] analyzed 13 Seckel families for linkage to SCKL1, but only one family was found linked to this locus. Also, a novel locus was identified and designated as SCKL3 in family 3 [10]. As a next step, our aim was to test our families in SCKL2. Initial studies at this locus were carried out by Uğur [44]. Some of the families had been analyzed with markers D18S869, D18S1107, D18S877 and D18S535. In this study, a total of 17 markers were employed in all families and the previous results were confirmed; thus, linkage was refined extensively at SCKL2.

A homozygous block with 5 consecutive microsatellite markers was detected in patient 15 at SCKL2. This candidate locus is just centromeric to the candidate gene *RBBP8* and may be extended to include this gene (Figure 5.1). In this patient also another homozygous region for three consecutive markers was detected, between markers D18S66 and D18S-S04. This 1.243 Mb region was centromeric to SCKL2. So, it is more plausible to consider the homozygous region inside the locus with primary attention and further analysis must be carried out by designing markers to test if this region extends to include *RBBP8* (Figure 5.1). An intragenic marker could be designed and employed for this gene to test whether it is included in the homozygous region. In families 9 and 12, a shared homozygous block of 1.531 Mb spanned by four markers at the telomeric side within SCKL2 was determined. These two families were small, and family 12 was informative for only two markers. In order to assess this region as a gene locus, further evidence must be obtained. To achieve this, additional markers could be designed inside the locus, and if these markers are found homozygous in patients, the probability that this region is the gene

locus would increase. Patient in family 11 was found to have a homozygous genotype between markers D18S-S04 and D18S847, but this region is also centromeric to SCKL2. Moreover, this region was continuously revised in NCBI database in the last updates and is still ambiguous. The genes inside this region could be evaluated only after this ambiguity is cleared. In summary, families 9, 11, 12 and 15 were found indicative of linkage to SCKL2, and the remaining families were excluded.

The region for which patient 15 is homozygous for 5 consecutive markers most likely harbors the gene for SCKL2. The locus could possibly include *RBBP8*, which has a role in DNA repair and thus a strong candidate for the disease in this family. Investigation of additional large Seckel families around the region where patients 9 and 11 are homozygous would reveal whether this region is really a SCKL locus, as the region does not contain any genes that could be implicated in the Seckel etiology. Patient 11 was found homozygous at the centromeric side of SCKL2, and this region does not contain any candidate gene either. Some of those homozygous regions could well be false positive results, since small homozygous blocks are not very rare in consanguineous families.

Families 3, 6, 7 and 12 were found indicative of linkage to SCKL3. The homozygous region shared by the families 3, 6 and 7 could be the most probable region. Families 6 and 12 shared another region homozygous for two markers in the telomeric region of SCKL3. Whether it is this region that harbors the disease gene might be clarified by analysis using additional markers within the region. The remaining eleven families were not indicative of linkage to SCKL3. It is proposed that genes acting in the same pathway would have similar consequences and lead to similar phenotypes.

In this study, investigation of the three SCKL loci was performed. Only one out of 15 families linked to SCKL1; 13 of the families were screened at this locus previously [43]. For SCKL2 the most positive results were obtained in family 15. The region must be evaluated and refined with additional markers, and then the smaller homozygous regions seen in families 9, 11 and 12 could be evaluated.

In the next step of the studies, important genes that directly interact with ATR were determined from genome databases and initial screening for linkage using nearby

microsatellite markers performed. After the analysis of *CHK2* locus with the nearby marker D22S689, patients in families 6, 7, 11 and 13 were found homozygous. Families 6 and 11 were noninformative and family 7 was informative for only one parent. This locus must be further evaluated with additional nearby markers. In the analysis of *CHK1*, two large families were found homozygous for D11S4464 but not for nearby markers, thus this locus was excluded with additional markers (D11S4158 and D11S968), as they were found heterozygous. Such false positive results complicate linkage analysis, but all should be evaluated carefully in order not to miss a real locus. RAD17 and RAD9 were also analyzed, but the markers employed were rather far from the gene loci. Those markers were considered relatively suitable, some genome wide scans are done with markers that are 25 Mb apart. However, these gene loci should also be evaluated with closer markers in order to confirm homozygosity in the families.

Our analysis was extended to include other genes implicated in DNA repair. Some of the genes were selected and tested for linkage. Among those, *DNAPK* was tested for linkage in all families. In family 11, a homozygous region spanning 6,3 Mb containing *DNAPK*, which has a role in DNA nonhomologous end joining repair (NHEJ), was detected. Also, allele 3 for marker D11S1110 for which the patient in this family was homozygous is a rare allele, increasing the probability that this allele came from a common ancestor carrying a putative mutation. These results initially pointed to a probable role of *DNAPK* in Seckel etiology, but as this family is small, further evidence is required to rule out a false positive. This region needs to be studied with additional markers. However, the gene with its eighty-seven exons is too large to screen for mutations. *DNA ligase IV* (*LIG4*), essential for NHEJ, was also tested for linkage. Seckel patients share many features with *LIG4* patients, so it would be necessary to screen our families for linkage to this gene. The basic difference of *LIG4* and Seckel patients is that pancytopenia is associated with the mutations in this gene. Although our families do not show pancytopenia, the first patient with a defect in DNA ligase did not have pancytopenia either and exhibited different clinical features than those patients identified by O'Driscoll [35]. So analysis of *LIG4* locus for linkage was performed. However, no linkage was observed, indicating that it is not a common locus for SCKL. For the genes tested, no common gene locus shared by several families was found. As these Seckel cases all have phenotypic differences, they probably have different genetic defects. Initial analyses for

these genes did not give promising positive linkage in all families. But they are important for guiding further comprehensive studies. Also, the large Seckel families with parental consanguinity used in this study are invaluable to search for candidate genes.

MNAT1, *RAD51L1* and *P2RX1* were analyzed for mutations in relevant families by direct sequencing, and SSCP, *NYD-SP20* were analyzed by direct sequencing, no sequence or pattern difference indicative of an association with the etiology of the diseases under investigation could be detected. However, failure to detect mutations in any of these genes does not absolutely exclude the gene as responsible for the relevant disease, since a causative mutation may lie deep in intronic regions not covered by the analyses. Such a mutation could affect proper functioning of an enhancer or silencer or create a novel splicing site that would disrupt proper splicing. cDNA studies are more efficient in detecting mutations affecting splicing, but this method has several drawbacks. Firstly, it is not easy to obtain patient tissue samples that express the genes and even determine which tissues express the gene. Secondly, the genes might be expressed in some of the developmental stages only. Promoter mutations are also mild mutations, so promoters should also be analyzed. In addition, patient 7, who was later found homozygous near *RAD51L1* locus, could be analyzed for mutations in the gene. Finally, the NCBI database could be searched for other possible candidate genes at the loci.

Initial linkage analysis to candidate DNA repair genes was carried out. The aim was to detect a gene to which many of the Seckel families were indicative of linkage. But such a common gene was not identified. In the future studies, the families which were found indicative of linkage to any of the repair genes could be analyzed further using additional markers to ascertain linkage. Later, mutation search can be performed in relevant patients. Such a study possibly will lead to the identification of novel loci. The genes that gave positive linkage for at least one family were *DNAPK*, *XRCC2*, *RAD52*, *RAD51C*, and *XRCC3*. However, some of them are rather large; for example *DNAPK* has 87 exons.

Some of the symptoms of the patients in family 8 resemble also Rett syndrome. This family for which 3 generations were available is currently being scanned genome-wide to identify the gene locus. After the gene locus is hopefully identified, the other families could then be analyzed for linkage to the locus.

Identification of new genes responsible for Seckel syndrome would shed light into the mechanisms of how genetic heterogeneity is achieved, and how defects in different genes cause similar phenotypes. As Seckel syndrome has a developmental background, identification of causative genes would enhance our understanding in development and in the contribution of genes mainly functioning in DNA repair to developmental processes.

In the azoospermia studies, the locus identified by Kılınç [43] was refined with additional markers. The homozygous region was delineated by D17S973. In the beginning *P2RX1* was assessed as a strong candidate. It spans the genomic region 4,016 – 4,026 Mb on chromosome 17, which includes the homozygous marker D17S1828. As the gene had been found essential for male reproduction in a mouse model, all of its 12 exons and 25-30 bp intronic regions were analyzed, but no mutation was found. Markers D17S1152 and D17S919, which were found heterozygous, interrupted the homozygosity region. Since no other candidate gene was present in the two homozygous regions, the heterozygous region was evaluated and *NYD-SP20* was noticed. As database errors occur often, these two heterozygous markers interrupting the homozygous region were hypothesized as misplaced in the genome database. It is very unlikely that a double-crossover occurred in this very small interrupting region, which could explain occurrence of a heterozygous region between two homozygous blocks. Therefore, the order of the markers was not convincing. Also, these markers were in different clones and the chromosome maps were built by combining these clones at their overlapping regions. During this process, some of the clones might have been misplaced.

The presence of over 200 genes inside this locus made it difficult to assess other candidate genes. Candidate genes other than those analyzed were: *P2RX5*, a member of the *P2RX* family, that shares homology with *P2RX1*, which is essential for male reproductive system; *SPAG7* encoding a sperm associated protein; *ODF4* encoding a sperm tail protein that plays an important role in sperm tail structure and sperm movement; *SOX15* encoding a protein involved in male gonad development and highly expressed in adult testis.

The maximum two-point lod score between the markers and azoospermia was 2.18, whereas seventeen-point lod score reached maximally to 3,259, demonstrating the power of multi-point analysis in autozygosity mapping. This value indicated significance of the

locus, and it is highly probable that the disease gene resides at the locus. The disease gene could possibly reside between markers D17S1308 and D17S695 where maximum lod score appeared. But no gene related to reproduction is present in this region. Also lod score was higher than 3 between markers D17S559 and D17S796. This also pointed out to the significance of this region. *SPAG7* is present in this region. *P2RX5* resides in the heterozygote region between these two homozygous blocks, between markers D17S1152 and D17S919, but might be misplaced in genome databases. Around genes *SOX15* and *ODF4* loci, the lod score was slightly less than 3, but these genes could not be excluded as the markers around the gene loci were not informative. The homozygosity regions on chromosome 17p13 have to be narrowed down in order to assess the strongest candidate genes in the locus.

REFERENCES

1. Majewski, F. and T. Goecke, "Studies of Microcephalic Primordial Dwarfism. I: Approach to a Delineation of the Seckel Syndrome", *American Journal of Medical Genetics*, Vol. 12, pp. 7-21, 1982.
2. Majewski, F. and J. Spranger, "A New Type of Primordial Dwarfism", *Monatsschr Kinderheilkd*, Vol. 124, pp. 499-503, 1976.
3. Majewski, F., M. Ranke and A. Schinzel, "Studies of Microcephalic Primordial Dwarfism. II: The Osteodysplastic Type II of Primordial Dwarfism", *American Journal of Medical Genetics*, Vol. 12, pp. 23-35, 1982a.
4. Majewski, F., M. Stoeckenius and H. Kemperdick, "Studies of Microcephalic Primordial Dwarfism. III: An Intrauterine Dwarf with Platyspondyly and Anomalies of Pelvis and Clavicles-Osteodysplastic Primordial Dwarfism Type III", *American Journal of Medical Genetics*, Vol. 12, pp. 37-42, 1982b.
5. Hersh, J. H., M. R. Joyce, J. Spranger, E. C. Goatley, R. S. Lachman, S. Bhatt and D. L. Rimoin, "Microcephalic Osteodysplastic Dysplasia", *American Journal of Medical Genetics*, Vol. 51, pp. 194-199, 1994.
6. Hayani, A., C. R. Suarez, Z. Molnar, M. Lebeau and J. Godwin "Acute Myeloid Leukaemia in a Patient With Seckel Syndrome", *Journal of Medical Genetics*, Vol. 31, No. 2, pp. 184-189, 1994.
7. Di Blasi S., M. Belvedere, S. Pintacuda, U. Di Blasi, N. Ferotti, S. Galante and D. Galante, "Seckel's Syndrome: A Case Report", *Journal of Medicine*, Vol. 24, No. 1, pp. 75-96, 1993.

8. Goodship, J., H. Gill, J. Carter, A. Jackson, M. Splitt and M. Wright, "Autozygosity Mapping of a Seckel Syndrome Locus to Chromosome 3q22. 1-q24", *American Journal of Human Genetics*, Vol. 67, No. 2, pp. 498-503, 2000.
9. Borglum, A. D., T. Balslev, A. Haagerup, N. Birkebaek, H. Binderup, T. A. Kruse and J. M. Hertz, "A New Locus for Seckel Syndrome on Chromosome 18p11.31-q11.2", *European Journal of Human Genetics*, Vol. 9, 2001.
10. Kilinc, M. O., V. N. Ninis, S. A. Ugur, B. Tuysuz, M. Seven, S. Balci, J. Goodship, and A. Tolun, "Is The Novel SCKL3 at 14q23 the Predominant Seckel Locus?", *European Journal of Human Genetics*, Vol. 11, No. 11, pp. 851-857, 2003.
11. Syrrou, M., I. Georgiou, M. Paschopoulos and D. Lolis, "Seckel Syndrome in a Family with Three Affected Children and Hematological Manifestations Associated with Chromosome Instability", *Genetic Counselling*, Vol. 6, No. 1, pp. 37-41, 1995.
12. Woods, C. G., M. Leversha and J. G. Rogers, "Severe Intrauterine Growth Retardation with Increased Mitomycin C Sensitivity: a Further Chromosome Breakage Syndrome", *Journal of Medical Genetics*, Vol. 32, No. 4, pp. 301-305, 1995.
13. Carney, J. P., R. S. Maser, H. Olivares, E. M. Davis, M. Le Beau, J. R. Yates, L. Hays, W. F. Morgan and J. H. Petrini, "The hMre11/hRad50 Protein Complex and Nijmegen Breakage Syndrome: Linkage of Double-Strand Break Repair to the Cellular DNA Damage Response", *Cell*, Vol. 93, No. 3, pp. 477-486, 1998.
14. O'Driscoll, M., V. L. Ruiz-Perez, C. G. Woods, P. A. Jeggo, J. A. Goodship, "A Splicing Mutation Affecting Expression of Ataxia-telangiectasia and Rad3-Related Protein (ATR) Results in Seckel Syndrome", *Nature Genetics*, Vol. 33, No. 4, pp. 497-501, 2003.

15. Yu, X. and R. Baer, "Nuclear Localization and Cell Cycle-Specific Expression of CtIP, a Protein that Associates with the BRCA1 Tumor Suppressor", *Journal of Biological Chemistry*, Vol. 275, pp.18541-18549, 2000.
16. D'Adamo, P., A. Menegon, C. Lo Nigro, M. Grasso, M. Gulisano, F. Tamanini, T. Bienvenu, A. K. Gedeon, B. Oostra, S. K. Wu, A. Tandon, F. Valtorta, W. E. Balch, J. Chelly and D. Toniolo, "Mutations in GDI1 are Responsible for X-Linked Non-Specific Mental Retardation", *Nature Genetics*, Vol 19, No. 2, pp. 134-139, 1998.
17. Adamczewski, J. P., M. Rossignol, J. P. Tassan, E. A. Nigg, V. Moncollin and J. M. Egly, "MAT1, cdk7 and Cyclin H Form a Kinase Complex Which is UV Light-Sensitive Upon Association with TFIIH", *EMBO Journal*, Vol. 15, No. 8, pp. 1877-1884, 1996.
18. Shinohara, A., H. Ogawa and T. Ogawa, "Rad51 Protein Involved in Repair and Recombination in *S.Cerevisiae* is a RecA-Like Protein", *Cell*, Vol. 69, No. 3, pp. 457-470, 1992.
19. Albala J. S., M. P. Thelen, C. Prange, W. Fan, M. Christensen, L. H. Thompson and G. Lennon, "Identification of a Novel Human RAD51 Homolog, RAD51B", *Genomics*, Vol. 46, No. 3, pp. 476-479, 1997.
20. Masson, J. Y., M. C. Tarsounas, A. Z. Stasiak, A. Stasiak, R. Shah, M. J. McIlwraith, F. E. Benson and S. C. West, "Identification and Purification of Two Distinct Complexes Containing the Five RAD51 Paralogues", *Genes and Development*, Vol. 15, pp. 3296-3307, 2001.
21. Meehan, W. J., R. S. Samant, J. E. Hopper, M. J. Carrozza, L. A. Shevde, J. L. Workman, K. A. Eckert, M. F. Verderame and D. R. Welch, "Breast Cancer Metastasis Suppressor 1 (BRMS1) Forms Complexes With Retinoblastoma-Binding Protein 1 (RBP1) and the mSin3 Histone Deacetylase Complex and

- Represses Transcription", *Journal of Biological Chemistry*, Vol. 279, No. 2, pp. 1562-1569, 2004.
22. Havre, P. A., M. C. Rice, M. Noe, and E. B. Kmiec, "The Human REC2/RAD51B Gene Acts as a DNA Damage Sensor by Inducing G1 Delay and Hypersensitivity to Ultraviolet Irradiation", *Cancer Research*, Vol. 58, No. 20, pp. 4733-4739, 1998.
 23. Lopez-Girona, A., "Serine-345 is Required for Rad3 Dependent Phosphorylation and Function of Checkpoint Kinase Chk1 in Fission Yeast", *The Proceedings of National Academy of Sciences*, Vol. 98, pp. 11289-11294, 2001.
 24. Goudelock, D. M., K. Jiang, E. Pereira, B. Russell and Y. Sanchez, "Regulatory Interactions Between the Checkpoint Kinase Chk1 and the Proteins of the DNA-Dependent Protein Kinase Complex", *Journal of Biological Chemistry*, Vol. 278, No. 32, pp. 29940-29947, 2003.
 25. Sorensen, C. S., R. G. Syljuasen, J. Falck, T. Schroeder, L. Ronnstrand, K. K. Khanna, B. B. Zhou, J. Bartek and J. Lukas, "Chk1 Regulates the S Phase Checkpoint by Coupling the Physiological Turnover and Ionizing Radiation-Induced Accelerated Proteolysis of Cdc25A", *Cancer Cell*, Vol. 3, No. 3, pp. 247-258, 2003.
 26. Bell, D. W., J. M. Varley, T. E. Szydlo, D. H. Kang, D. C. Wahrer, K. E. Shannon, M. Lubratovich, S. J. Verselis, K. J. Isselbacher, J. F. Fraumeni, J. M. Birch, F. P. Li, J. E. Garber and D. A. Haber, "Heterozygous Germ Line hCHK2 Mutations in Li-Fraumeni Syndrome", *Science*, Vol. 286, No. 5449, pp. 2528-2531, 1999.
 27. Matsuoka, S., M. Huang and S. J. Elledge, "Linkage of ATM to Cell Cycle Regulation by the Chk2 Protein Kinase", *Science*, Vol. 282, pp. 1893-1897, 1998.
 28. Chehab, N. H., A. Malikzay, M. Appel and T. D. Halazonetis, "Chk2/hCds1 Functions as a DNA Damage Checkpoint in G-1 by Stabilizing p53", *Genes and Development*, Vol. 14, pp. 278-288, 2000.

29. Bao, S., R. S. Tibbetts, K. M. Brumbaugh, Y. Fang, D. A. Richardson, A. Ali, S. M. Chen, R. T. Abraham and X.-F. Wang, "ATR/ATM-Mediated Phosphorylation of Human Rad17 is Required for Genotoxic Stress Responses", *Nature*, Vol. 411, pp. 969-974, 2001.
30. Volkmer, E. and L. M. Karnitz, "Human Homologs of *Schizosaccharomyces Pombe* Rad1, Hus1, and Rad9 Form a DNA Damage-Responsive Protein Complex", *Journal of Biological Chemistry*, Vol. 274, pp. 567-570, 1999.
31. Greer, D.A., B. D. Besley, K. B. Kennedy and S. Davey, "hRad9 Rapidly Binds DNA Containing Double-Strand Breaks and is Required for Damage-Dependent Topoisomerase II Beta Binding Protein 1 Focus Formation", *Cancer Research*, Vol. 63, No. 16, pp. 4829-4835, 2003.
32. Anderson, C. W. and S. P. Lees-Miller, "The Nuclear Serine/Threonine Protein Kinase DNA-PK", *Critical Reviews in Eukaryotic Gene Expression*, Vol. 2, pp. 283-314, 1992.
33. Hartley, K. O., D. Gell, G. C. M. Smith, H. Zhang, N. Divecha, M. A. Connelly, A. Admon, S. P. Lees-Miller, C. W. Anderson and S. P. Jackson, "DNA-Dependent Protein Kinase Catalytic Subunit: a Relative of Phosphatidylinositol 3-Kinase and the Ataxia Telangiectasia Gene Product", *Cell*, Vol. 82, pp. 849-856, 1995.
34. Robins, P. and T. Lindahl, "DNA Ligase IV from HeLa Cell Nuclei", *Journal of Biological Chemistry*, Vol. 271, pp. 24257-24261, 1996.
35. O'Driscoll, M., K. M. Cerosaletti, P.-M. Girard, Y. Dai, M. Stumm, B. Kysela, B. Hirsch, A. Gennery, S. E. Palmer, J. Seidel, R. A. Gatti, R. Varon, M. A. Oettinger, H. Neitzel, P. A. Jeggo and P. Concannon, "DNA Ligase IV Mutations Identified in Patients Exhibiting Developmental Delay and Immunodeficiency", *Molecular Cell*, Vol. 8, pp. 1175-1185, 2001.

36. Rave-Harel, N., I. Madgar, R. Goshen, M. Nissim-Rafinia, A. Ziadni, A. Rahat, O. Chiba, Y. M. Kalman, C. Brautbar, D. Levinson, A. Augarten, E. Kerem and B. Kerem, "CFTR Haplotype Analysis Reveals Genetic Heterogeneity in the Etiology of Congenital Bilateral Aplasia of the Vas Deferens", *American Journal of Human Genetics*, Vol. 56, pp. 1359-1366, 1995.
37. Chillon, M., T. Casals, B. Mercier, L. Bassas, W. Lissens, S. Silber, M.-C. Romey, J. Ruiz-Romero, C. Verlingue, M. Claustres, V. Nunes, C. Ferec and X. Estivill, "Mutations in the Cystic Fibrosis Gene in Patients with Congenital Absence of the Vas Deferens", *New England Journal of Medicine*, Vol. 332, pp. 1475-1480, 1995.
38. Tiepolo, L. and O. Zuffardi, "Localization of Factors Controlling Spermatogenesis in the Nonfluorescent Portion of the Human Y Chromosome Long Arm", *Human Genetics*, Vol. 34, pp. 119-124, 1976.
39. Chandley, A. C., J. R. Gosden, T. B. Hargreave, G. Spowart, R. M. Speed and S. McBeath, "Deleted Yq in the Sterile Son of a Man with a Satellited Y Chromosome (Yqs)", *Journal of Medical Genetics*, Vol. 26, pp. 145-153, 1989.
40. Vogt, P., A. C. Chandley, T. B. Hargreave, R. Keil, K. Ma and A. Sharkey, "Microdeletions in Interval 6 of the Y Chromosome of Males with Idiopathic Sterility Point to Disruption of AZF, a Human Spermatogenesis Gene", *Human Genetics*, Vol. 89, pp. 491-496, 1992.
41. Mulryan, K., D. P. Gillerman, C. J. Lewis, C. Vial, B. J. Leckle, A. L. Cobb, J. E. Brown, E. C. Conley, G. Buell, C. A. Pritchard and R. J. Evans, "Reduced vas Deferens Contraction and Male Fertility in Mice Lacking P2X1 Receptors", *Nature*, Vol. 403, pp. 86-89, 2000.
42. Cheng, L. J., J. M. Li, J. Chen, Y. H. Ge, Z. R. Yu, D. S. Han, Z. M. Zhou and J. H. Sha, "NYD-SP16, a Novel Gene Associated with Spermatogenesis of Human Testis", *Biological Reproduction*, Vol. 68, pp. 190-198, 2003.

43. Kılınç, O., *Genotype-Phenotype Correlation in Cystic Fibrosis and Gene Localization in Seckel Syndrome and Pectus Carunatum*, Ph.D. Thesis, Boğaziçi University, 2001.
44. Uğur, S., *Locus and Gene Analysis in Five Inherited Disorders*, M.S. Thesis, Boğaziçi University, 2002.
45. Kavaslar, G. N., S. Önengüt, O. Derman, A. Kaya, and A. Tolun, "The Novel Genetic Disorder Microhydranencephaly Maps to Chromosome 16p13.3-12.1", *American Journal of Human Genetics*, Vol. 66, pp. 1705-1709, 2000.

