

# Single Nucleotide Polymorphisms associated with the Intronic *Cis* Regulatory Regions of *PAX7*: A Potential Linkage to Increased Tumorigenesis of Rhabdomyosarcoma elucidated via *In Silico* Biology and Pyrosequencing™

\*Maika G. Mitchell<sup>1,2</sup>, Diane Tabarini<sup>2</sup> Melanie Ziman<sup>1</sup>

<sup>1</sup> School of Exercise, Biomedical and Health Science, Edith Cowan University, Perth, W. Australia 6027

<sup>2</sup> Sloan Kettering Institute (Memorial Sloan Kettering Cancer Center), New York City, New York 10021, [blackmam@mskcc.org](mailto:blackmam@mskcc.org)

**Abstract:** Motivation. Single nucleotide polymorphisms (SNPs) exist as a type of genetic variation that can be either pathogenic or non pathogenic based on their influence on phenotype. At the date of the writing of this paper, there are 402 single nucleotide polymorphisms associated with intronic regions of human *PAX7*, which is found on chromosome one. Of these 75 are present in the intronic gene region of *PAX7* associated with alveolar rhabdomyosarcoma (ARMS) mainly found in the 3 prime regions of introns 5,6,7 and 8. The aim of this research is to identify those SNPs most likely to affect *PAX7* mediated tumorigenicity, particularly for alveolar rhabdomyosarcoma with which *PAX7* is associated. The research elucidated in this paper also examines the efficacy and ease of using *in silico* biological methods *prior* to the commencement of bench work to increase the success rate of SNP detection and allele quantification. Methods. *In Silico* Biology and Pyrosequencing™ are employed to confirm results obtained by conventional PCR gel electrophoresis and Sanger sequencing. Results. *In Silico* Biology and Pyrosequencing technologies are shown to be faster, more cost-efficient than conventional PCR and Sanger sequencing and provide excellent adjunct techniques to identify alterations in gene sequence. For *PAX7*, seventy-five intronic SNPs were identified to be within close vicinity (50 - 100 nucleotides) of intronic *cis* regulatory elements and some of these remain within the *PAX7* gene region present in the *PAX7-FOX1* fusion gene associated with alveolar rhabdomyosarcoma. Of these seventy-five, five were found to be different in alveolar rhabdomyosarcoma patients, both by conventional PCR, Sanger sequencing and Pyrosequencing. Changes due to LOH were revealed by the current genotyping results obtained by *in silico* means from the NCBI website and thus may provide an indication of altered sequences that contribute substantially to increased tumorigenicity. Moreover, six SNPs were within a conserved sequence in intron 5 & 8 of *PAX7* that is highly conserved in homologous *Pax7* genes of several animal phyla as well as in the exon/intron 23 regions of *NF-1* and intron 7 of *PAX3*. [Nature and Science. 2006;4(4):6-20].

**Keywords:** Single Nucleotide Polymorphisms; *PAX7*; Rhabdomyosarcoma; pyrosequencing, sequencing; ERMS, ARMS.

**Abbreviations and notations:** ARMS, alveolar rhabdomyosarcoma; ERMS, embryonal rhabdomyosarcoma cell lines; CNP, Copy Number Polymorphism; AQ, Allele Quantification; PSQ, Pyrosequencing; RD, Rhabdomyosarcoma; SNP, Single Nucleotide Polymorphism; TSS, transcription start site

## 1. INTRODUCTION

Single nucleotide polymorphisms (SNPs) exist as a type of genetic variation that can be either pathogenic or non-pathogenic based on their influence on phenotype. Up to the date of writing this paper, 402 single nucleotide polymorphisms have been identified within the intragenic region of the human *PAX7* gene locus, on chromosome 1. In a previous study [Barr et al.,

2002], exonic sequences were examined for the presence of SNPs that could be associated with *PAX7* mediated pathogenesis. In this paper we have searched the intronic chromosomal sequences of *PAX7* for the presence of SNPs that may be associated with increased tumorigenesis. While it is well known that SNPs within the exonic regions of a gene may change the protein structure and affect the functionality of the resultant protein, SNPs within intronic regions are less commonly associated with changes in functionality. However changes to intronic sequences may affect gene

expression levels or mRNA stability and/or gene splicing which may also lead to increased pathogenesis and tumorigenesis. The research elucidated in this paper identifies intronic SNPs that may be associated with *PAX7* mediated tumorigenicity in the childhood tumour alveolar rhabdomyosarcoma. Here we also examine the efficacy and ease of using *in silico* biological methods *prior* to the commencement of bench work to increase the success rate of SNP detection and allele quantification. Our findings show that mutations do indeed occur within the immediate vicinity (50 – 100 nucleotides) of the given *cis-acting* elements found in previous research (Mitchell et al., 2006).

## 2. MATERIALS AND METHODS

### *In Silico* Investigation

The introns of *PAX7* were scrutinized with five programs (three online, two offline and saved on PCs) which denote *cis* regulatory regions and/or transcription start sites (TSS) in submitted sequences. In our recent paper (Mitchell et al., 2006) we identified several *cis* regulatory elements in intronic regions of *Pax7*. The focus of the research for this paper is centered on identification of polymorphisms in the intronic regions of *Pax7* that may be associated with these *cis* regulatory elements. To do this the National Center for Biotechnology Information (NCBI) database for SNPs, incorporated into NCBI's Entrez system was queried to identify SNPs. Similarly Entrez databases, PubMed and GenBank were also queried. The Boolean advanced query allows one to use limits to restrict your search by search field, chromosome, and percentage of heterozygosity. To date (June 2006), we have identified four hundred-two SNPs associated with *PAX7* introns. After refining the search further based on the location of the SNPs (exon vs. intron) within the NCBI SNP database, seventy-five SNPs were identified as being associated with *PAX7* intronic regions and rhabdomyosarcoma, that is they remain within the *PAX7* region of the *PAX7-FOX1* fusion gene that is associated with alveolar rhabdomyosarcoma and are present within the intronic regions of *PAX7* found aberrantly expressed in some embryonic rhabdomyosarcomas. Furthermore, the seventy-five SNPs of introns 5, 7 and 8 were chosen for further investigation since the SNP locale statistics show that these three intronic regions had the highest concentration of single point polymorphisms. The list of seventy-five SNPs were analyzed relative to the results obtained by *in silico* methods to identify those polymorphisms close to previously identified *cis* regulatory regions ( Mitchell et al., 2006). The finalized list contained thirty-six SNPs associated with

the intronic regions of *PAX7* and rhabdomyosarcoma and are included in the transcript variants of *PAX7*.

### Tumor Samples and Control Samples

DNA from five ARMS samples was analyzed; all five samples were from tumours previously identified as having the *PAX7-FKHR* (1:13) translocation associated with ARMS (Tang et al., 1999) and were obtained from the Biomaxx, Inc. One of the DNA samples was isolated from an ARMS cell line (CW9019), and four were DNA isolated from patient tumour samples. In addition DNA was isolated from ERMS cell line ATCC#: CCL 136. In addition, DNA was isolated from buccal swabs of five normal volunteers.

### DNA Isolation

DNA from cell lines and tumour samples was isolated using Genra Autopure LS; the volunteer samples were extracted with the BuccalAmp™ DNA extraction kit, QuickExtract™ DNA extraction solution, and Catch-All™ sample collection swabs (EPICENTRE® Biotechnologies <http://www.epibio.com>) were used for the DNA isolation of the five normal samples submitted by volunteers.

### Primer & Assay Design

The primer design for Pyrosequencing was performed by utilizing the Biotage AB Assay Design Software version 1.0.6. Thirty to sixty base pairs 5' and 3' of the target polymorphisms were entered into the Assay Design program to identify the best forward, reverse, and sequencing primers. The automated program has as its output a list of scored (100=best, 60 or below passed to poor) primers for the target SNP(s).

### PCR Analysis / Gel Electrophoresis

Because of the limited quantity of DNA available on the ARMS cases for this study, a DNA dilution-based PCR assay was used for allele quantification and SNP detection. The following DNA concentrations (serial dilutions) were examined for the five cases; 2, 1, 0.5, 0.25, 0.125 ng/ul. One ul of each dilution was used as the DNA template for the PCR reactions. PCR was performed in a volume of 25 ul with the presence of 100 uM of each dNTP, 10 mM Tris-HCl (pH 8.3), 2 mM magnesium chloride, 50 mM potassium chloride, 0.75 unit of Taq DNA polymerase, and 15 pmoles of each custom primer. The following reaction conditions were employed: denaturation at 96°C for 5 min,

followed by 60 cycles each of 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C, with a final extension of 5 min at 72°C. Sixty cycles were used due to the small size of the fragment and the low concentration of the isolated patient DNA. Subsequently, aliquots of the reaction were resolved on 2% / 4% Invitrogen™ commercial pre-cast agarose E-gel, and the products were visualized by Sybr-Safe Green staining.

### Pyrosequencing

Twenty microliters of the biotinylated PCR products was immobilized on streptavidin-coated magnetic beads (10 µl of Dynabeads M-280 streptavidin solution [Dyna, Oslo, Norway]) in 25 µl of 2x binding-washing buffer II (10 mM Tris-HCl, 2 M NaCl, 1 mM EDTA, and 0.1% Tween 20 [pH 7.6]) at 65°C for 5 min in a shaking mixer (1,400 rpm). The PCR product-Dynabeads complexes were captured with a PSQ 96 Sample Prep Tool (Pyrosequencing AB, Uppsala, Sweden) and transferred to a 96-well PSQ 96 SQA microtiter plate containing 0.5 M NaOH (50 µl per well), and single-stranded DNA was obtained through incubation for 4 min.

### DNA Sanger Sequencing Big Dye Terminator Reaction

Sanger Sequencing has been around since 1977 [Rosen et al., 2000; Murphy et al., 2005]. Many advances have been made since then in terms of technology and the race to completely sequence the human genome. It was quickly realized that there are many SNPs and tandem repeats throughout the genome. Sanger sequencing is a very expensive, time consuming process requiring hours to review data and create contiguous traces using Phred and Phrap. The results of the combined procedure (Sanger and Pyrosequencing) are seen below in figures 5. After the Big Dye Terminator reaction of the samples, ethanol precipitation was used to remove excess dye. Twenty microliters of Hi-Di formamide was used and then the samples were denatured for four minutes at 98°C. Then, the sequencing samples were allowed to cool for 5 minutes. Run time of the samples in the 3730XL DNA Sequencer (Applied Biosystems©) was 1 hour. Unfortunately, without performing PCR and fragment analysis, allele quantification is not possible from standard Sanger sequencing.

### 3. RESULTS OF *IN SILICO* BIOLOGY

*In silico* methods were used to scan intronic regions of *PAX7* for SNPs (this paper) and *cis* regulatory regions (Mitchell et al., 2006). SNPs (three) located within or

close to *cis* elements were identified for further investigations *in vitro* (Figs. 1-3).

***In Silico* PCR Amplification** (<http://insilico.ehu.es/>) & (<http://genome.ucsc.edu/cgi-bin/hgPcr>)

*In-Silico* PCR amplification was performed on the three SNPs identified by *in silico* methodology described above. *In-Silico* PCR amplification provides the user a systematic approach to prescreen the custom-designed primers, before purchase for the given template. It also elucidates the prediction of probable PCR products and search of potential mismatching location of the specified primer(s). The output of these programs allow for quick and easy changes to primer design or template choices. This process was employed for every custom-designed primer created by the PSQ software prior to purchase. The two websites compare the user's template and primers.

### 4. RESULTS OF *IN VITRO* STUDIES BASED ON RESULTS OF *IN SILICO* METHODS

#### Standard PCR Analysis/ Gel electrophoresis

Standard PCR assays performed on all seven samples revealed amplification of the three selected SNPs found in the intronic regions of the *PAX7* gene (Figure 4). The gene-specific fragment was visualized at all dilution levels for all these cases. Figure 4 depicts positive PCR amplification for the control sample using primers 23, 24, and 40.

#### Pyrosequencing™ for allele quantification and confirmation of SNP detection

Pyrosequencing™ by Biotage©, although first initiated in 1993 (Sanger et al., 1977; Nyren et al., 1993) is based on the Sanger sequencing by synthesis method. Data from the Pyrosequencing method is a quantitative measure of each detected nucleotide and is used for measuring the amounts of alleles. This property allows the quantification of heterozygosity, multi-copy genes, pooled DNA samples, and mixed genotypes in heterogeneous samples (i.e. tumor and normal cells). In essence the user can get results from one Pyrosequencing plate that would equal four separate processes (RT-PCR, Big Dye Terminator cycle sequencing, and fragment analysis). Due to the small size of the fragment, twenty microliters of PCR product was used in the pyrosequencing reactions performed here. The biotinylated primers identified by *in silico* methods shown above were used to process the samples. Results are shown below. The total processing time of the samples was eight minutes, confirming the dramatic decrease in processing time over conventional Sanger sequencing.

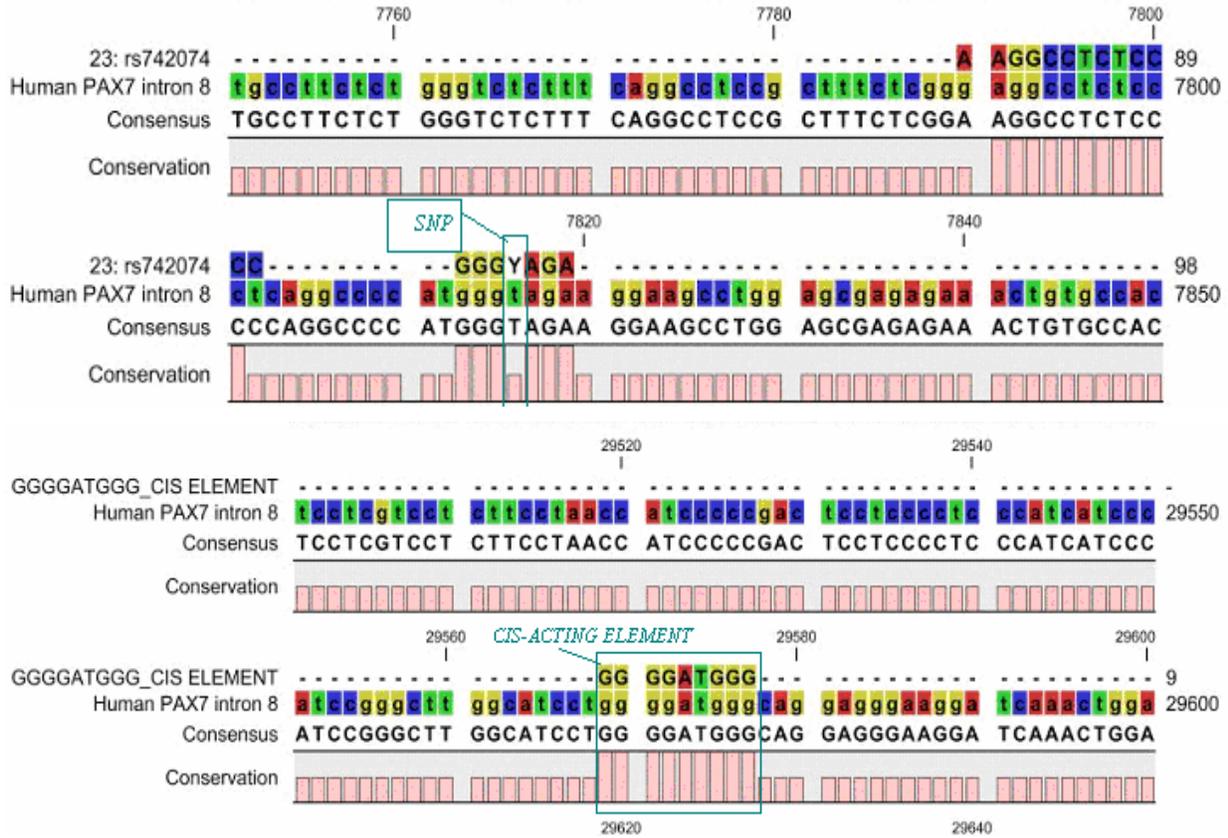


Figure 1. Cis element **GGGGATGGG** indicated in teal; (Mitchell and Ziman 2006) found in Human PAX7 intron 8 located near SNP rs742074 (Y = IUPAC code for (C/T)).

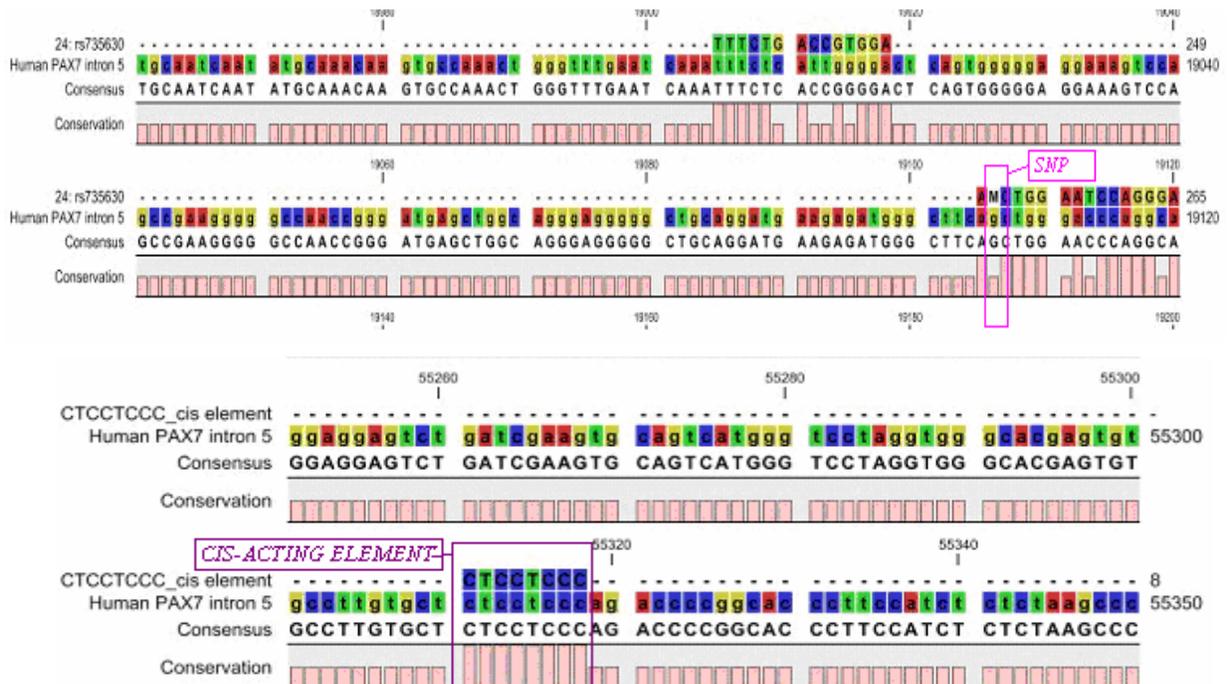
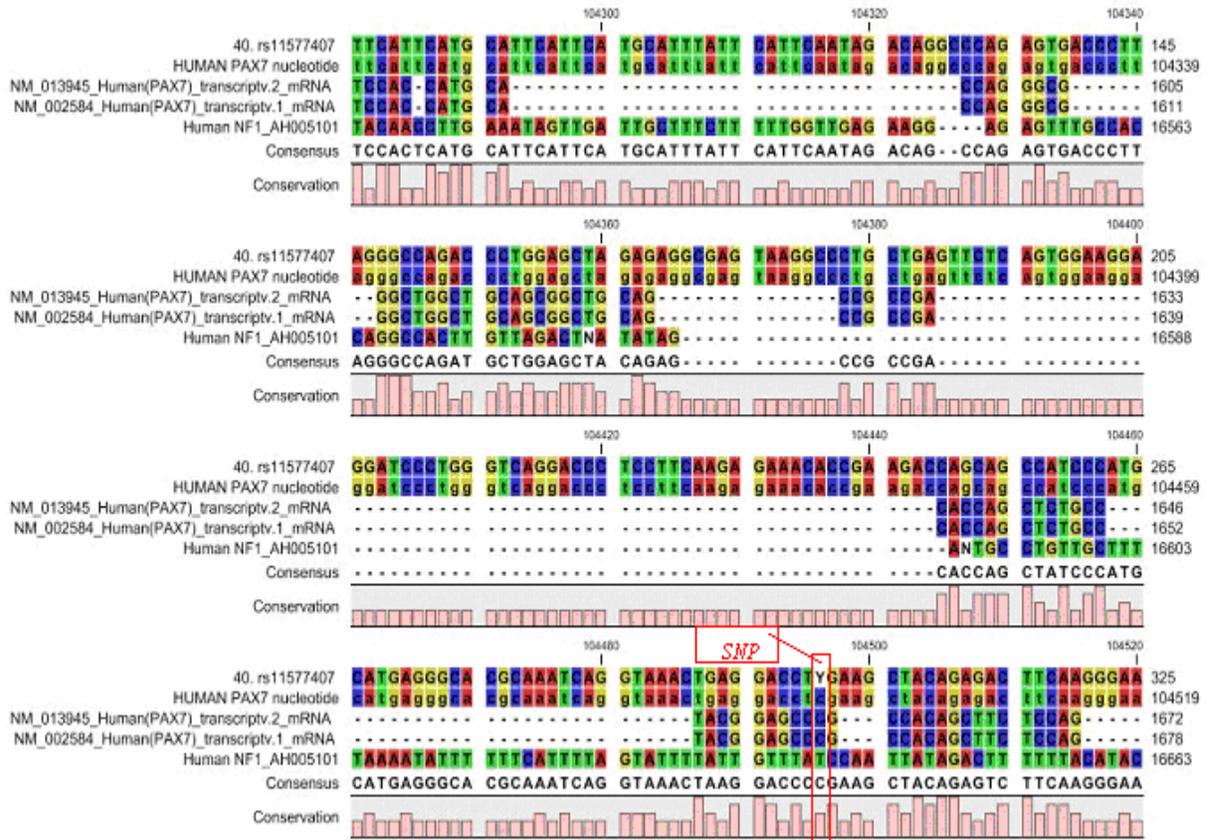


Figure 2. Cis element **CTCCTCCC** indicated in pink (Mitchell and Ziman 2006) found in Human PAX7 intron 8 located before SNP rs735630 (M = IUPAC code for C/A).



**Figure 3. Multiple alignment of PAX7 transcript variants ( NM\_013945 and NM\_002584 respectively), Human NF1, and the single nucleotide polymorphism found in the 3' end of Human PAX7 intron 8 located before SNP rs11577407 (Y = IUPAC code for C/T) indicated in orange.**

**Cis –acting element: GGGGATGGG**

**Primer 23: rs742074** [*Homo sapiens*] Intron 8 chromosome position: chr1:18802796-18802962

Genotype Detail

T: 0.955

C: 0.045

Primer	Sequence
Forward PCR Primer	TAATCCACCCCTCTCAAGAATG
Reverse PCR Primer	TGAGATTTGCCTTCAGATAAAACC
Sequencing Primer	ACCAGGAAAAAAGTAAAT

**Cis –acting element: CTCCTCCC**

**Primer 24: rs735630** [*Homo sapiens*] Intron 5 chromosome position: chr1:18726993-18727052

Genotype Detail

A: 0.950

C: 0.050

Primer	Sequence
Forward PCR Primer	CGCAGCCCTTGTTTCTGA
Reverse PCR Primer	CAGGGGCCAGAAGCTGGT
Sequencing Primer	CTTGTTTCTGACCGTG

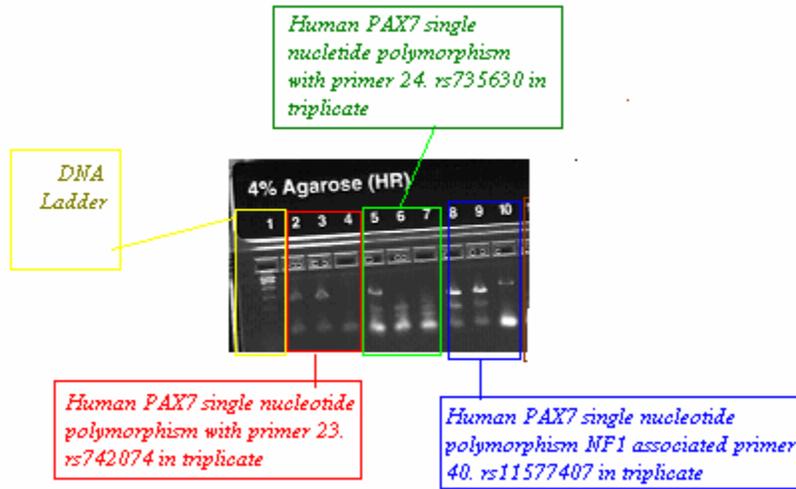
**NF1 associated Primer 40: rs11577407** [*Homo sapiens*] Intron 8 chromosome position: chr1:18807266-18807391

Primer	Sequence
Forward PCR Primer	ATGAGGGCACGCAAATCAG
Reverse PCR Primer	ACCCAGAGACACGAGCAC
Sequencing Primer	AATCAGGTAACTGAGGAC

Genotype Detail

C:0.998

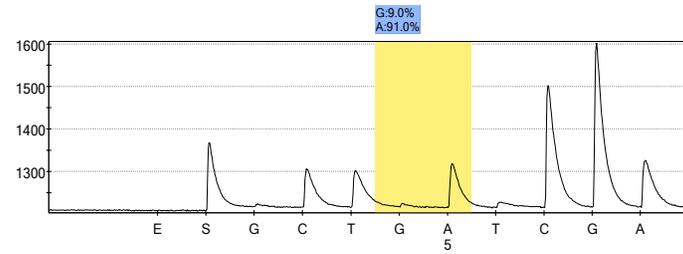
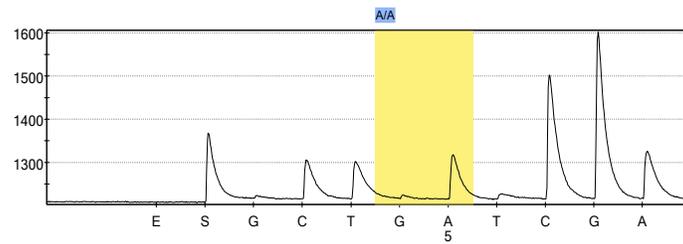
T:0.002



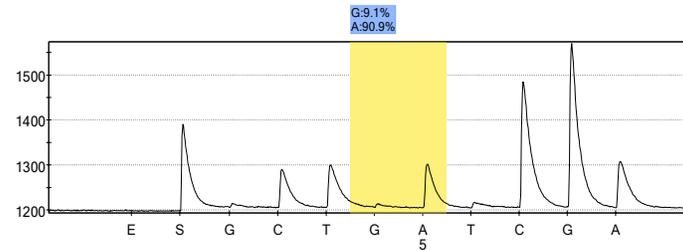
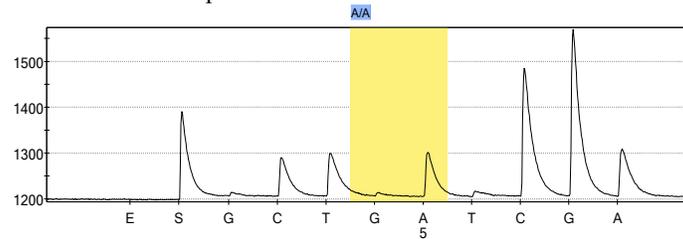
**Figure 4.** Invitrogen™ 4% precast Agarose Gel electrophoresis showing serial dilution for DNA isolated from volunteer buccal swab. Sybr-Safe staining resolved on Bio-Rad UV Gel Doc 2000. Gel electrophoresis of DNA isolated from ATCC #136 ERMS Cell line (not shown) , ARMS patients(1-5; (not shown)) and control sample (shown below) showed similar results with the biotinylated primers 23,24 and 40.

Figure 5 shows the results of the Biotage© Pyrosequencing™ PSQ HS 96A & Sanger sequencing for Primers 23, 24 & 40 for ARMS, ERMS and Control Samples (Figure 5) .

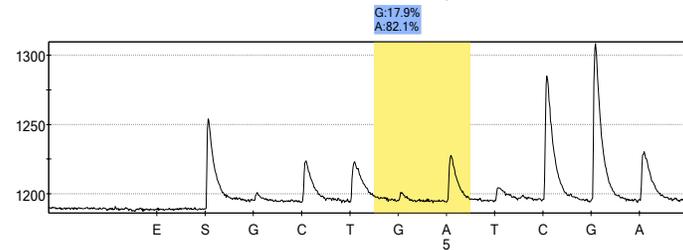
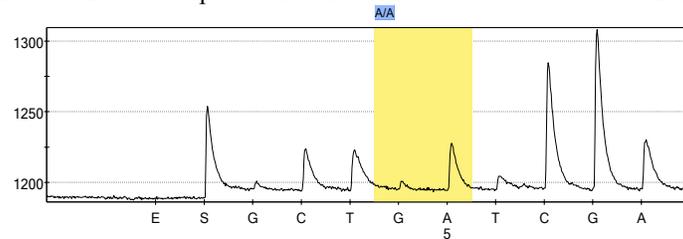
23: rs742074 Sample: ARMS PATIENT 1 : G: 9.0% / A: 91.0%



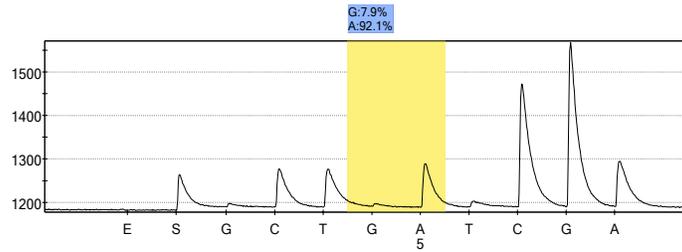
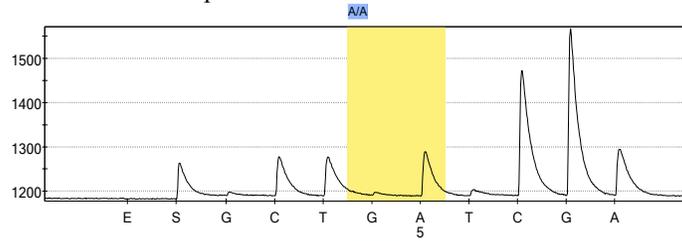
23: rs742074 Sample: ARMS PATIENT 2 : G: 9.1% / A: 90.9%



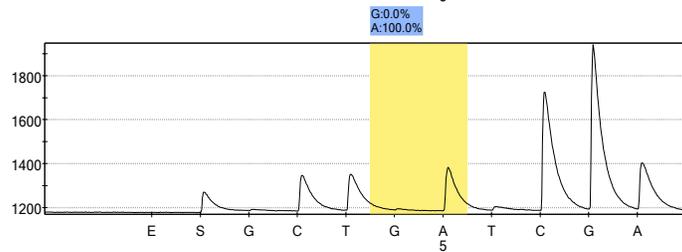
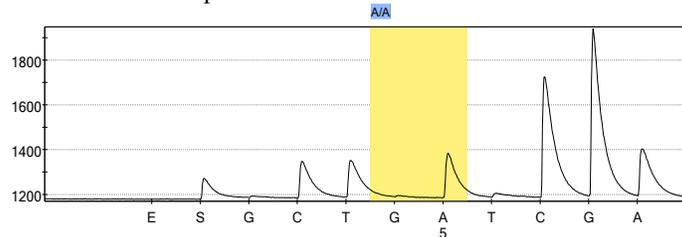
23: rs742074 Sample: ARMS PATIENT 3 : G: 17.9% / A: 82.1%



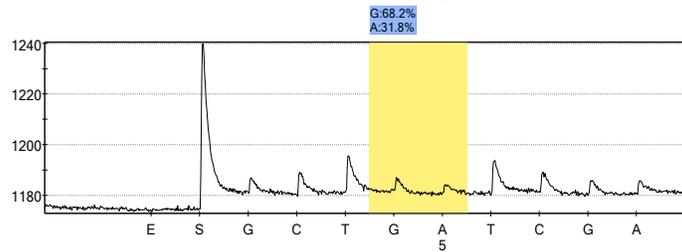
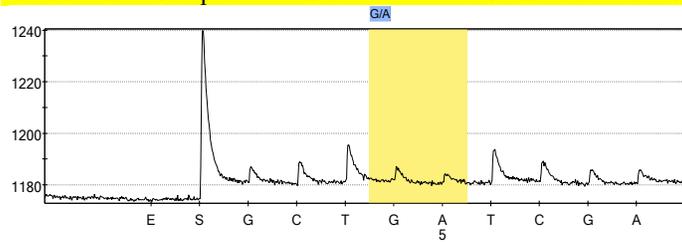
23: rs742074 Sample: ARMS PATIENT 4 : G: 7.9% / A: 92.1%



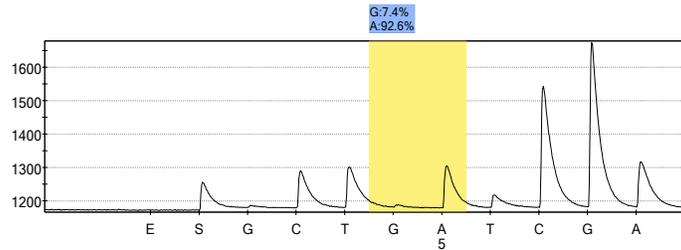
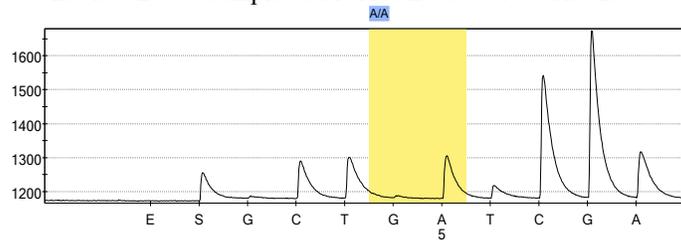
23: rs742074 Sample: ARMS PATIENT 5: G: 0.0% / A: 100.0%



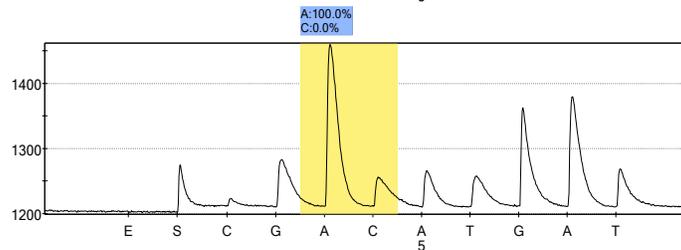
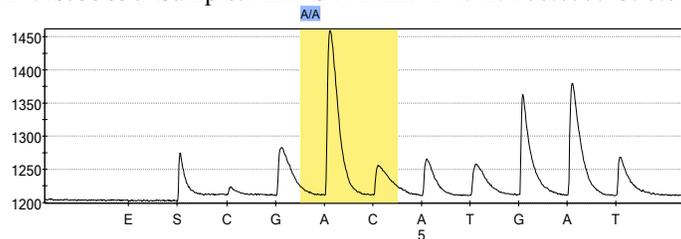
23: rs742074 Sample: ERMS CELL LINE : G: 68.2% / A: 31.8%



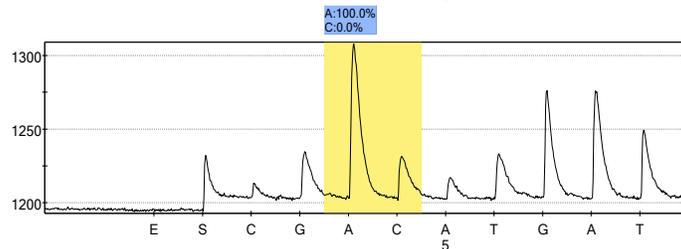
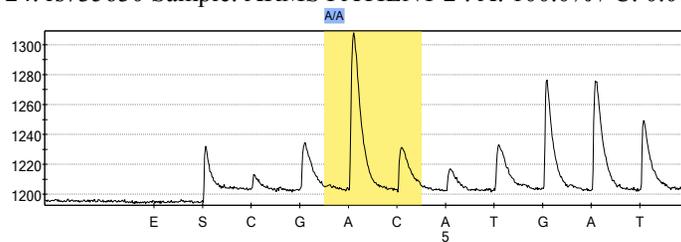
23: rs742074 Sample: CONTROL : G: 7.4% / A: 92.6%



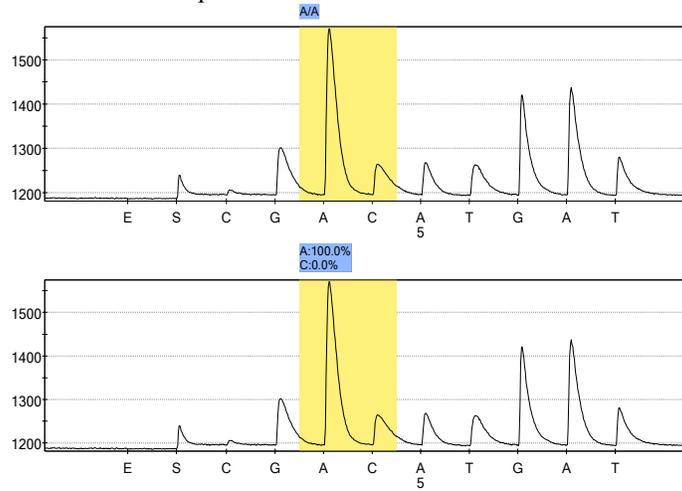
24: rs735630 Sample: ARMS PATIENT 1: A: 100.0% / C: 0.0%



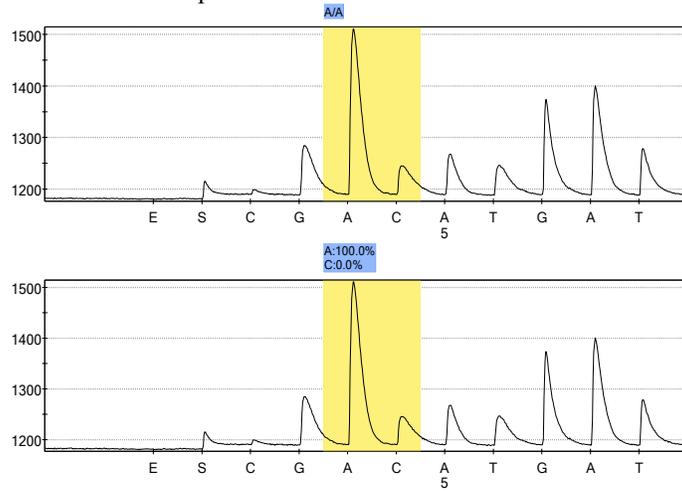
24: rs735630 Sample: ARMS PATIENT 2 : A: 100.0% / C: 0.0%



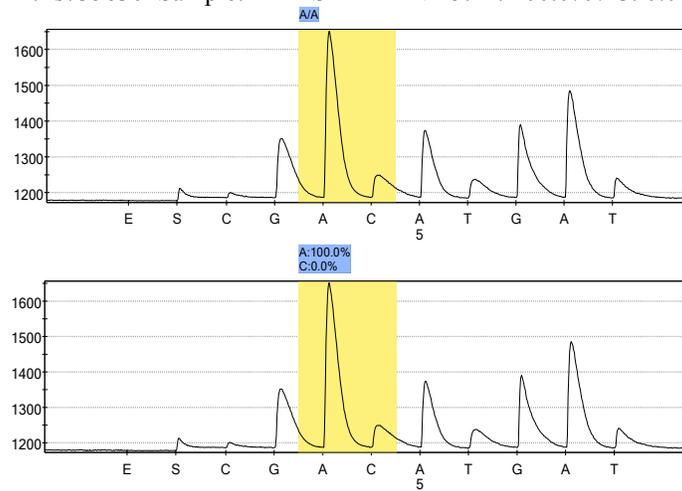
24: rs735630 Sample: ARMS PATIENT 3 : A: 100.0% / C: 0.0%



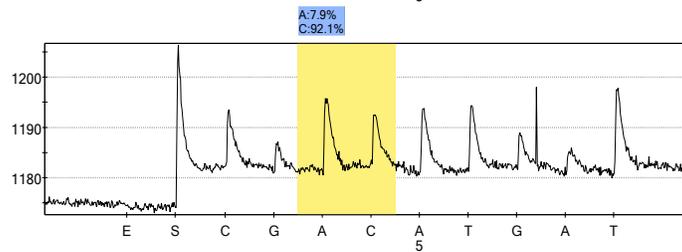
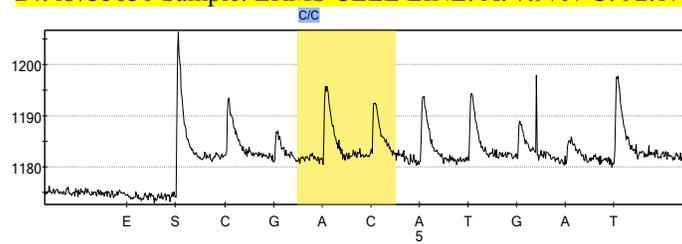
24: rs735630 Sample: ARMS PATIENT 4 : A: 100.0% / C: 0.0%



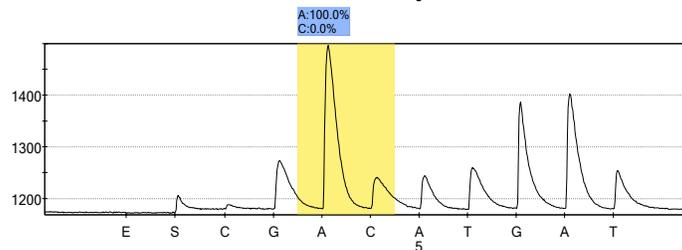
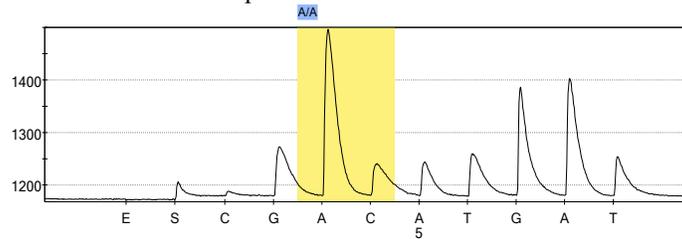
24: rs735630 Sample: ARMS PATIENT 5 : A: 100.0% / C: 0.0%



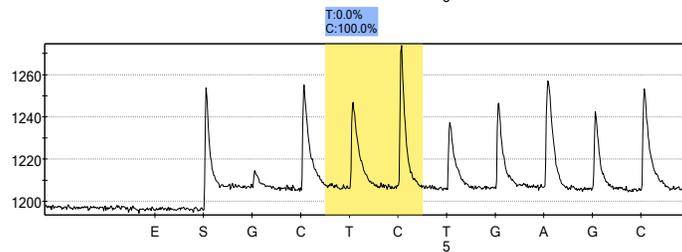
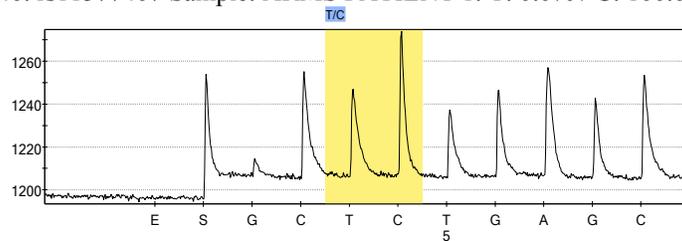
24: rs735630 Sample: ERMS CELL LINE: A: 7.9% / C: 92.1%



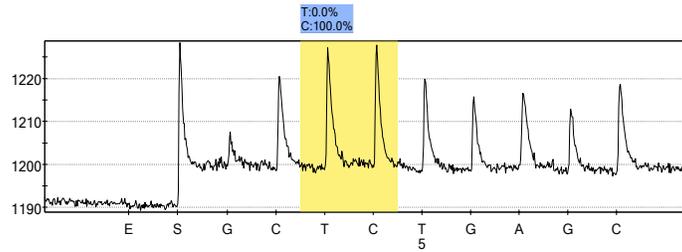
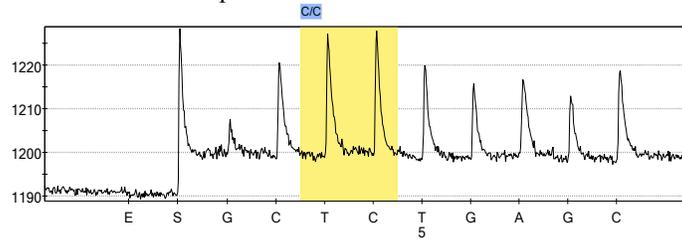
24: rs735630 Sample: CONTROL: A: 100.0% / C: 0.0%



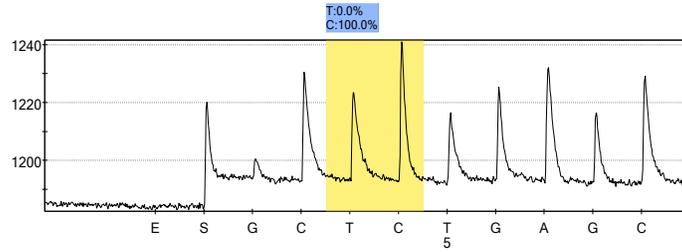
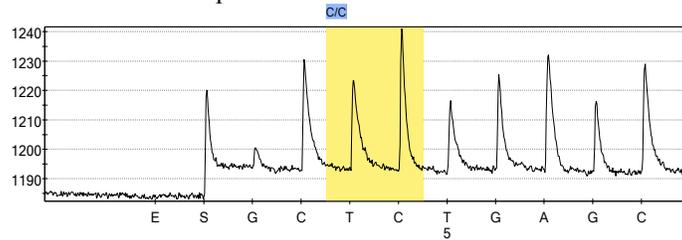
40. rs11577407 Sample: ARMS PATIENT 1: T: 0.0% / C: 100.0%



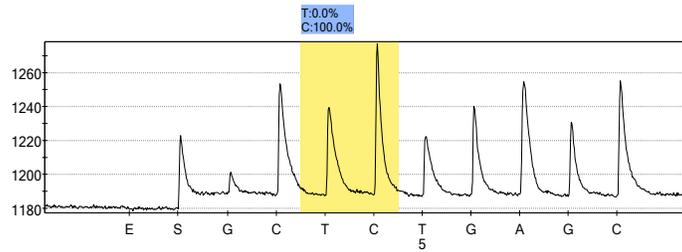
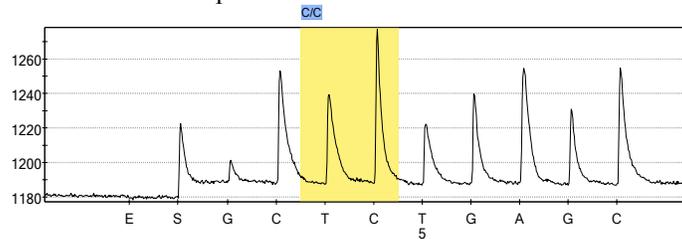
40. rs11577407 Sample: ARMS PATIENT 2: T: 0.0% / C: 100.0%



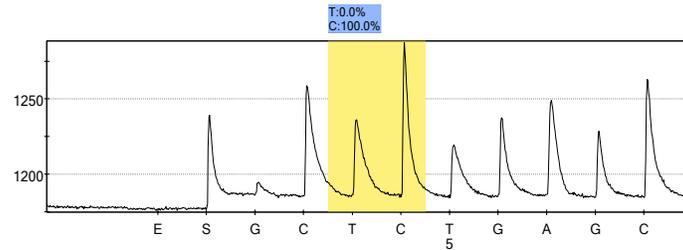
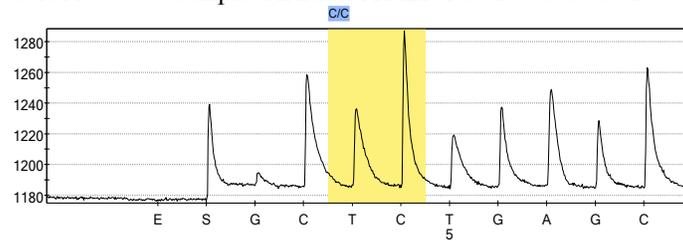
40. rs11577407 Sample: ARMS PATIENT 3: T: 0.0% / C: 100.0%



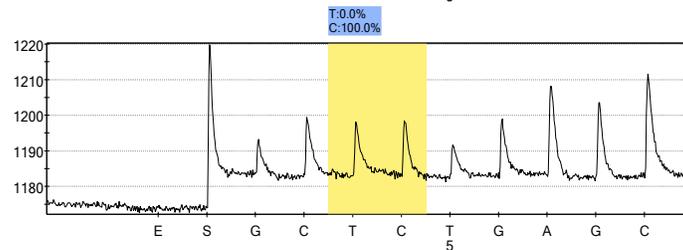
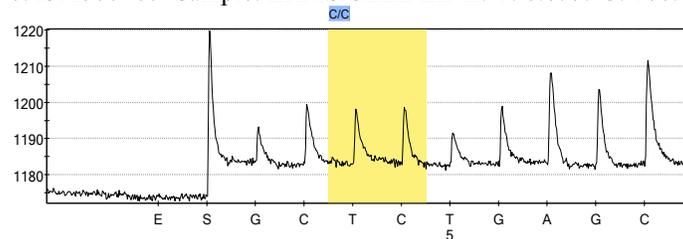
40. rs11577407 Sample: ARMS PATIENT 4: T: 0.0% / C: 100.0%



40. rs11577407 Sample: ARMS PATIENT 5: T: 0.0% / C: 100.0%



40. rs11577407 Sample: ERMS CELL LINE: T: 0.0% / C: 100.0%



40. rs11577407 Sample: CONTROL: T: 0.0% / C: 100.0%

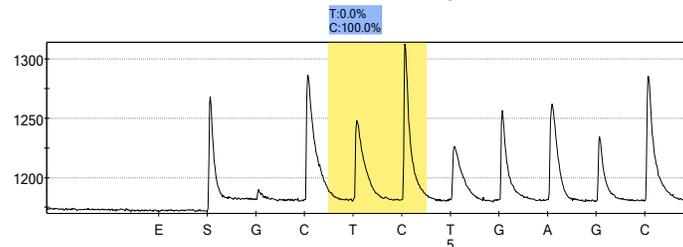
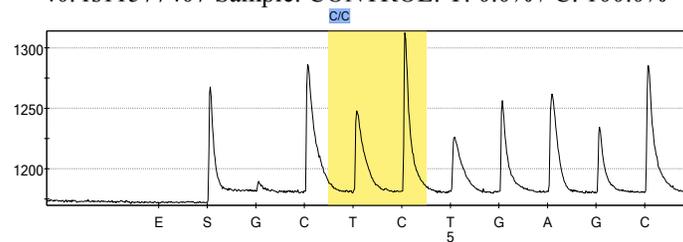


Figure 5. The results of the Biotage© Pyrosequencing™ PSQ HS 96A & Sanger sequencing for Primers 23, 24 & 40 for ARMS, ERMS and Control Samples.

## 5. DISCUSSION

### Comparison of *In silico* PCR Analysis and Pyrosequencing to conventional *in vitro* PCR Analysis/ Gel electrophoresis

Any new process introduced into the scientific field must undergo rigorous testing. It takes many publications to initiate the establishment of new methods. The methods applied in this paper while not novel technology, are not common place in research laboratories. *In silico* biology PCR analysis provides one with a prescreening of the functionality of the primers chosen. It does not ensure that your PCR reaction will work. Those issues are resolved at the bench where the constraints on success range from template quality to varying concentrations of primers to thermocycler parameters. The programs do however give some assurance to the scientist during the ordering process of primers and reagents. The programs also decrease the time spent on trial and error of the selection of primer. What normally would have taken two weeks to perform was reduced to one week. *In silico* PCR also provided a troubleshooting method of where the primer (s) would anneal their  $T_M$ , and possible creation of secondary structures within the oligo itself or with the given template. The size of the fragment from the use of paired primers also helps the scientist to select the proper concentration of agarose gel to use (0.8%, 1.0%, 2.0% and so on...) and the number of cycles for needed to amplify the target sequence.

The results observed when comparing the two processes were very helpful in highlighting the cost aspects associated with development of the genotyping assay itself. The time taken from ordering the oligos to Sanger sequencing was three days. Two additional days were required to process and screen for the Pyrosequencing to confirmation the results. This procedure may become a new way to reduce the time downstream of all SNP detection, discovery and allele quantification algorithms.

### Pyrosequencing Results

Results presented here show that in DNA isolated from the ERMS cell line, two variants in genotype and allele quantification were identified, one for primer 23 A→G in intron 8 region, and for primer 24 C→A in intron 5 of Human PAX7. This may be an anomaly of cell culture yet no changes were identified in the ARMS cell line. The pyrosequencing method used for studying this gene was selected to minimize the risk of missing a mutation and for accurate allele quantification. Therefore it is possible that with this

new technology and short processing time, has revealed previously unidentified changes in allele frequency associated with ERMS. The importance of the findings presented here should be viewed in the light of the significance of finding changes in allele frequency associated with any form of cancer. If LOH is detected it may well indicate disease severity or susceptibility to the disease. It is therefore important to investigate all SNPs for their association with tumorigenicity allowing physicians to refine the care and treatment of patients with these diseases. Further research is required to assess the number of cases of ERMS that show similar allele changes. Our studies highlight the ease and speed with which these studies may be undertaken.

Interestingly there were no changes in SNP frequency observed in the ARMS samples relative to the expected allele frequency at the selected SNPs. Since only a few SNP sites were investigated in this paper, additional SNP analysis at other identified sites may reveal significant changes in allele frequency and LOH in ARMS patients. Our findings do however highlight the differences in the ARMS and ERMS sarcomas; ARMS is associated with a chromosomal translocation that produces a *PAX7-FOX1* fusion gene, whereas ERMS is associated with increased expression of *PAX7* (Robson *et al.*, 2006). The increased expression of Pax7 in ERMS samples may be linked to the changes in allele frequency or LOH at the sites identified. Further research is required to confirm this association.

### Acknowledgment

We would like to thank the Memorial Sloan Kettering Institute DNA Sequencing Core Facility and Memorial Sloan Kettering Cancer Center's Diagnostic Molecular Laboratory for their support and assistance for the *in vitro* portions of this project.

**Received:** October 1, 2006

### REFERENCES

- [1] Barr, Frederic G, Qualman, Stephen J, Macris, Michele H, Melnyk, Natalya, Lawlor, Elizabeth R, Strzelecki, Donna M, Triche, Timothy J, Bridge, Julia A, Sorensen, Poul HB. Genetic Heterogeneity in the Alveolar Rhabdomyosarcoma Subset without Typical Gene Fusions *Cancer Res* 2002;62:4704-4710.
- [2] Kathleen M. Murphy, Karin D. Berg, James R. Eshleman, Sequencing of Genomic DNA by Combined Amplification and Cycle Sequencing Reaction, *Clin Chem* 2005;51:35-39.
- [3] Steve Rozen and Helen J. Skaletsky Primer3 on the WWW for general users and for biologist

- programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, 2000:365-386.
- [4] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **1977**;74:5463-5467.
- [5] Nyren P, Pettersson B, Uhlen M. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. *Anal. Biochem.* 1993;208:171-175.
- [6] Ronaghi M. Pyrosequencing Sheds Light on DNA Sequencing. *Genome Research*. 2001;11(1):3-11.
- [7] Ina Vandembroucke, Tom Callens, Anne De Paepe, Ludwine Messiaen. Complex splicing pattern generates great diversity in human NF1 transcripts. *BMC Genomics*. 2002;3:13-23.
- [8] Gutmann DH, Geist RT, Rose K, Wright DE. Expression of two new protein isoforms of the neurofibromatosis type 1 gene product, neurofibromin, in muscle tissues. *Dev Dyn*. **1995**;202:302-311.
- [9] Gutmann DH, Zhang Y, Hirbe A. Developmental regulation of a neuron-specific neurofibromatosis 1 isoform. *Ann Neurol*. **1999**;46:777-782.
- [10] Rozen S, Skaletsky H., Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press. **2000**:365-386.
- [11] Bikandi J, San Millán R, Rementeria A, Garaizar, J. *In silico* analysis of complete bacterial genomes: PCR, AFLP-PCR, and endonuclease restriction. *Bioinformatics*. **2004**;20(5):798-9.
- [12] Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977;74:5463-5467.
- [13] Tang ED, Nuñez G, Barr FG, Guan KL. Negative regulation of the forkhead transcription factor FKHR by Akt. *J Biol Chem*. 1999;274(24):16741-16746.
- [14] Mitchell, Maika G, Ziman, Melanie. An In Silico Investigation into the Discovery of Novel Cis-acting Elements within the Intronic Regions of Human PAX7. *Nature & Science* 2006;4(3)12:69-85.
- [15] Wong, Kwok, Tsang Y. et al., Allelic imbalance analysis by high density single nucleotide polymorphic allele (SNP) array with whole genome amplified DNA. *Nucl Acids Research*. 2004;32:1-8.
- [16] Robson E, He S, Eccles M. A Panorama of PAX Genes in Cancer and Development. *Nat Rev Cancer*. 2006;6(1):52-62.