AV.

Institute of Neurology
(Queen Square)
University College London

Rare mutation on PCM1 causes susceptibility to schizophrenia

Msc in Clinical Neuroscience Research Project 2008

Hui-Minn Chan

Bsc (Hons) in Psychology 2006
University of York, UK

UCL Institute of Neurology
The National Hospital for Neurology and Neurosurgery
Queen Square
London WC1N 3BG

&

Molecular Psychiatry Laboratory
Department of Mental Health Sciences
Royal Free and University College London Medical School
Windeyer Institute of Medical Sciences
London W1T 4JF

MSc Clinical Neuroscience
2007/08

Submitted as partial fulfilment of the requirements for the MSc in Clinical Neuroscience, University College London

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

Being on the MSc course has been an incredible journey, and my sojourn at the Molecular Psychiatry lab in the Windeyer Institute of Medical Sciences proved a befitting finale as this chapter of my life draws to a close. I have many to thank, hence writing this particular bit of my project gives me the greatest pleasure of all.

I would like to start by thanking Professor Hugh Gurling who has been a brilliant supervisor and an amazing mentor in every aspect. I am forever grateful to Dr Andy McQuillin, for his tireless efforts in coaching a "genetics idiot", for his generosity in sharing his knowledge, and especially for his unwavering optimism.

I also want to thank Radhika Kandaswamy and Ana Pereira, PhD extraordinaire of the lab, for mentoring me and cheering for me every step of the way. Many thanks also to Vinay Puri, Nick Bass, and members of the lab who have helped me in many little ways even if it was just to give encouragement.

I am also indebted to my tutor Dr Anthony Pullen, the best there ever was and is, for his guidance throughout the course. Advice and kindness from Dr Caroline Selai and members of the Education Unit will always be cherished. Special thanks to Ms Daniela Warr for coordinating my leave of absence due to extenuating circumstances, and to Ms June Smalley for helping out with the printing.

Lastly and most importantly, I want to thank my parents for their unfaltering support and understanding all along, without which I will never be who I am today.

.....

Statement of contribution:

Study design: Dr Andy McQuillin, Hui-Minn Chan

Data collection: Hui-Minn Chan, Dr Andy McQuillin

(Special mention: Radhika Kandaswamy, Ana Pereira)

Data analysis: Hui-Minn Chan, Dr Andy McQuillin

Writing up: Hui-Minn Chan, Professor Hugh Gurling, Dr Andy McQuillin

Table of contents

51 pages in total; 5 Tables and 18 Figures / Diagrams

Abbreviations

PCM1	Pericentriolar-material 1
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
ssDNA	Single-stranded DNA
PCR	Polymerase chain reaction
HRM	High Resolution Melting
DISC1	Disrupted-In-Schizophrenia-1
DTNBP1	Dysbindin
NRG1	Neuregulin1
RGS4	Regulator of G-protein signalling 4
BBS	Bardet-Biedel Syndrome
LOD	Logarithm of odds (to the base of 10)
SNP	Single Nucleotide Polymorphism
SADS-L	Schedule for Affective Disorders and Schizophrenia – Lifetime
HAP1	Huntington Associated Protein

Abstract

Schizophrenia is a major mental illness that affects 0.85% of the population and has been shown to have a strong genetic component. Previous family linkage studies have shown that there is a susceptibility locus present in the chromosome 8p21-22 region. Subsequently allelic association studies have shown that the *pericentriolar-material 1* (PCM1) gene is involved in genetic susceptibility to schizophrenia. Several potential disease mutations in PCM1 have been found. One of these was a rare G→T mutation on PCM1 exon 24. This mutation was found to co-segregate with schizophrenia in a single family. The objective of this research project was to replicate this finding in a sample of 613 cases and 624 control DNA samples. One of the methods used was a novel real-time polymerase chain reaction (PCR) High Resolution Melting point assay on DNA from cases and controls. Artificially-designed positive control templates incorporating the single base change were created. After genotyping, all potential mutants were sequenced. The mutation was discovered in at least 2 of the 613 cases while still awaiting sequencing results for the 624 controls. The two new mutations are being followed up in the families of the probands to check whether the mutation is found in other affected relatives. It is proposed that mutations involving PCM1 may cause a subtype of schizophrenia, and that rare mutations are important in predisposing towards schizophrenia.

1. Introduction

An estimated 0.85% of people worldwide are afflicted by schizophrenia (Saha et al., 2005), a devastating disease widely regarded as the cancer of mental illness. The disease burden is extremely high because it usually manifests during late adolescence or early adulthood. Evidence from family, twin, and adoption studies provide strong support for genetic influences (Sullivan, 2008), although efforts to uncover the genetics basis of schizophrenia has been complicated by its genetic heterogeneity and phenotype complexity. About 100 genes have recently been implicated in genome-wide association studies (Gurling, unpublished data).

Before the genome-wide association studies had been conducted, only a few genes including Disrupted-In-Schizophrenia-1 (DISC1), Dysbindin (DTNBP1), neuregulin1 (NRG1), and Regulator of G-protein signalling 4 (RGS4) had been considered possible candidates in causing susceptibility to schizophrenia (Harrison & Weinberger, 2005; Sullivan, 2008). PCM1 [OMIM 600299], shown in Figure 3, has been shown to interact directly with DISC1 at the protein level (Miyoshi et al., 2004). PCM1 and DISC1 have both been shown to be crucial for neurite outgrowth, neuronal migration, dendritic arborisation, and hence proper development of the cerebral cortex (Kamiya et al., 2005). PCM1 also interacts with other proteins such as those involved in Bardet-Biedel Syndrome (BBS), an inherited disorder characterised by cognitive impairment, renal dysfunction, and obesity (Kim et al., 2004). Interestingly, BBS sufferers seem to be at a higher risk of developing schizophrenia (Ansley et al., 2003).

In itself, PCM1 is an important component of centriolar satellites, targets several proteins to the centrosome, and plays an important role in regulating microtubular dynamics (Kubo et al., 1999; Kubo & Tsukita, 2003). Previously, 4 independent linkage analysis studies have indicated the 8p21-22 region in schizophrenia with LOD (logarithm of odds) scores above 3.00, or 1000:1 in favour of genetic linkage (Blouin et al., 1998; Brzustowicz et al., 1999; Gurling et al., 2001; Suarez et al., 2006). Another linkage study reported a LOD score between 2.0 and 3.0 (Kendler et al., 1996). A mutation on exon 28 of PCM1 substituting amino acid threonine with isoleucine (SNP rs370429), likely to change the structure and function of PCM1, has been found to be associated with schizophrenia (Gurling et al., 2006). This finding however was not replicated (Sawa et al., 2008, Personal communication) but small sample size was cited as an explanation. In Gurling et al.'s (2006) study, schizophrenics with an inherited PCM1 abnormality were compared with those without and were discovered to have significantly reduced grey matter volume in the orbitofrontal cortices. The researchers concluded that it was possible that mutations in PCM1 cause a subtype of schizophrenia biased towards affective and behavioural features.

Recently, a group of researchers exploring centrosomal dysfunction as an underlying risk for psychiatric illnesses (Sawa et al., 2008) discovered a G→T base pair change on exon 24 of PCM1. This heterozygous mutation (named 4057G→T), seemingly rare as it was found only in 1 out of 32 cases, introduces a premature stop codon (E1353X) which may produce a truncated version of the protein or lead to nonsense-mediated decay (Sawa et al., 2008). The mutation was also found in 2 members of the proband's family who suffer from schizophrenia-spectrum

psychosis (Figure 2). Like Gurling et al. (2006), Sawa and colleagues (2008) have proposed that variations on the PCM1 gene may be linked to a subtype of schizophrenia. They speculate that it may be deficit schizophrenia, a subgroup of schizophrenia with enduring and idiopathic negative symptoms as the hallmark of disease (Carpenter et al., 1988). Suggestions of this nature may well prove to be timely as schizophrenia is a disease with various clinical presentations including psychotic symptoms and personality disorders. The complexity in phenotype may cloud interpretations of genetics studies. Indeed, Kirkpatrick et al. (2000) have suggested that the deficit-non-deficit dichotomy may be useful as a phenotype in genetic linkage studies.

Given that PCM1 is strongly-implicated in the aetiology of schizophrenia and that the 4057G→T mutation seemed to be highly penetrant at least according to Sawa et al.'s (2008) findings, it was decided that a replication should be carried out to detect other cases with the introduced stop codon on exon 24.

1.1. Uncovering the 4057G→T mutation

A DNA sample carrying the 4057G→T mutation was unavailable; thus positive DNA templates of the wild type (GG) and mutants (GT, TT) had to be produced. The equipment utilised, the Lightcycler® 480 Instrument (Roche), uses real-time polymerase chain reaction (PCR) fluorescent detection of genetic variation with a combination of different fluorophores and amplification approaches. The current research made use of two of these modules – namely high resolution melting and endpoint genotyping – prior to sequencing. The basic principles underlying the two

methods used in uncovering the single nucleotide polymorphism (SNP) of interest are discussed below.

1.1.1. High resolution melting (HRM) analysis

Briefly, the rationale behind the HRM analysis is as follow: DNA with a heterozygote SNP undergoes PCR and becomes available in high copy number. PCR is performed in the presence of an intercalating and saturating dye which binds only to double-stranded DNA (dsDNA). Therefore, at the beginning of a melting experiment, the initial fluorescence is high, and as the temperature is raised, the dsDNA dissociates into single-stranded DNA (ssDNA) and fluorescence diminishes. Each melting curve observed is characteristic of a particular DNA sample, and the shapes of the melting curves for heterozygotes and homozygotes are rather different. Heterozygotes are distinguished by their lower melting temperature because the heteroduplices they form are less stable than homoduplices formed between the wild type or variant homozygotes. The principles are illustrated in Figures 4 to 6.

1.1.2. Endpoint genotyping with the KASPar chemistry

Endpoint genotyping is a method for allele discrimination using allele-specific amplification for each allele of the SNP (Figure 7). The chemistry used for endpoint genotyping in this project was the relatively new KBiosciences Competitive Allele-Specific PCR SNP genotyping system (KASPar) which relies on the ability of specially-modified KTaq polymerase to detect polymorphisms. Two oligonucleotides are designed specific to each allele of the SNP and a common reverse primer is used. The allele-specific primers are designed with two different 5' tails (to incorporate

the allele-specific dyes) and two different 3' allele-specific bases. The KTaq polymerase distinguishes between alleles as it is unable to extend beyond mismatched bases at the 3' end and thus yields the fluorescence from the two allele-specific reporter dyes (fluorescein, or FAM, and VIC) included in the reaction.

In short, PCM1 appears a highly-plausible candidate in causing susceptibility to schizophrenia based on linkage studies and due to its interactions with key genes like DISC1. Its importance in ensuring normal growth of the cerebral cortex has also been highlighted. This research project aims to replicate a single base pair mutation which may produce an abnormal protein or lead to loss of protein function.

2. Methods

2.1. Schizophrenia and control samples

The UCL schizophrenia cohort of 613 case and 624 control samples were collected from various collaborating hospitals and clinics from London and South England. Diagnoses were made using the SADS-L (Schedule for Affective Disorders and Schizophrenia – Lifetime) diagnostic system (Spitzer & Endicott, 1977) by trained psychiatrists. The "supernormal" control subjects were also screened to ensure no psychiatric disorder was present in the control or first degree relatives in order to maximise the chances of finding a genetic difference between case and control groups. Genomic DNA was extracted from frozen whole blood samples using a standard cell lysis, proteinase K digestion, phenol/chloroform, ethanol precipitation method. All DNA samples were quantified with picogreen (Molecular Probes) by fluorimetry. A full description of the schizophrenia and control samples can be found in Datta et al., (submitted).

2.2. Artificial DNA template synthesis

Artificial DNA templates corresponding to 121 base pairs in the region of exon 24 incorporating the 4057G→T mutation were synthesised using 2 common forward PCR primers (E1353X_TEMP_F1 and E1353X_PCR1) and 4 reverse PCR primers – 2 allele-specific (E1353X_TEMPG_R1 and E1353X_TEMPT_R1) and 2 common (E1353X_TEMP_R2 and E1353X_PCR_R1) – ordered from Sigma Genosys. These primers have all undergone high performance liquid chromatography (HPLC) purification. The sequences of the primers making up the G- and T-allele containing

templates are shown in Table 1. The 121 base-pair template sequence prepared from the overlapping forward and reverse primers is as below:

E1353X G (normal / abnormal)
GAACAGACATTCAGCCCAGACTGAAGAGCCTGTTCAAGCAAAAGTATTCAGCAGAAAGAA
TCAT***g/t***AGCAACTGGAAAAAATAATAAAATGTAATAGGTCTACAGAAATATCTTCAGGTAT
G

Following unsuccessful template synthesis using BIOTAQ Red DNA polymerase (Bioline) – see section 2.4 on cloning and sequencing– template preparation was carried out over 3 stages of PCR with a proofreading Platinum® *Pfx* DNA Polymerase kit (Invitrogen) instead. The forward and reverse primers were added at different stages of the process and all primers were prepared at 10pmol/µL (refer Table 2). The products from the previous stage served as templates for the subsequent stage, shown in Figure 8.

All reactions utilised a PCR amplification protocol of 25 cycles with a denaturation step at 95°C for 15 seconds (s), an annealing step at 55°C for 30s, an extension step at 68°C for 30s, and a synthesis step at 72°C for 7 minutes (m). Upon termination of cycling the reactions were maintained at 8°C. Gel electrophoresis with AquaPōr™ HR GTAC™, a 4.5% high resolution agarose gel made following the given protocol, was used to analyse PCR products from Stages 1 – 3. HyperLadder V (Bioline) was used as the reference, and when BioTaq Red was not used 2µL Loading Buffer (Bioline) was added to the PCR products.

2.3. High-resolution melting (HRM) analysis of artificial templates

The PCR products from Stage 3 were used as DNA templates to make up positive controls of homozygous GG and TT as well as heterozygous GT. Water was included

as negative control. E1353X_PCR_F1 and E1353X_PCR_R1 were used as the forward and reverse primers respectively along with LightCycler® 480 High Resolution Melting Master (Version March 2007, Roche) reagents. The High Resolution Melting Dye included in the kit binds to the amplified PCR products and yields the fluorescence signals used to detect melting curves specific to each genotype. The optimal magnesium chloride ($MgCl_2$; co-factor for DNA polymerase) concentration for the primer pair used was determined via serial dilutions of 1.0mM to 3.5mM (in 0.5 increments). Out of the many HRM assays performed, the optimal $MgCl_2$ concentration varied between 2.0 and 3.0mM. The final 5µL (4 of Master Mix, 1 of template DNA) reaction was amplified using the recommended PCR protocol on the LightCycler® 480 Instrument (Roche):

95°C for 10m		Pre-incubation
95°C for 10s 65°C - 53°C for 10s 72°C for 15s	}	45 cycles
95°C for 1m 40°C for 1m 65°C for 1s 95°C continuous	}	High resolution melting
40°C for 10s		Cooling

2.3.1. Troubleshooting

The first HRM assay was unsuccessful and it was thought that the ratio of the G to T allele making up the heterozygote was inaccurate and that PCR was over-efficient due to an extremely high concentration of the templates. Hence, subsequent attempts included a 1 in 1000 dilution of the templates, different combinations of

G:T ratio for the heterozygote, as well as a higher volume of Master Mix (9µL) in an attempt to increase PCR efficiency. Concurrently, the artificial templates were also cloned and sequenced (see following section for relevant methods) to ensure that the SNP of interest was indeed incorporated into the templates. Unfortunately this did not produce the desired observation and was consequently abandoned (see below).

It was decided that HRM should proceed despite the inability to get a clear sequence of the artificial templates. Fresh Stage 3 templates were created and the 5µL Master Mix and template DNA reactions optimised and subjected to HRM on 384-well plates. The assay was also set up to include 1µL of 12.5ng/µL genomic DNA (gDNA) from 3 random controls in place of the templates. Improvisations included using SensiMix HRM™ (Quantace) with EvaGreen® dye instead of LightCycler® 480 HRM Master reagents and amending the touchdown protocol for the PCR annealing stage to 55°C on the premise that this was the condition used during template synthesis stages.

2.4. Cloning and sequencing of artificial DNA template produced

2.4.1. Cloning

To ensure that the homozygous artificial templates contained the G and T allele respectively, preparations were made to clone and sequence them. 1mL of 100µg/mL of ampicillin was added to 1L of dissolved Luria Bertani (LB) agar obtained from the Windeyer stores. The addition of ampicillin would lead to the selective survival of competent cells with the plasmid insert as the pDrive vector (Qiagen) confers resistance to the antibiotic. The agar was then poured into plates

and upon solidification, 50µg of 40mg/mL of 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside (X-Gal) and 100µL of 100mM Isopropyl b-D-1-thiogalactopyranoside (IPTG) were plated for blue/white screening, with white recombinant colonies containing the vector insert (ie along with the DNA template).

To ensure that the DNA to be cloned was free from PCR debris or contamination for optimal cloning results, it was cleaned using *micro*CLEAN (Microzone Ltd) following the protocol supplied. A Qiagen® cloning kit with 2X Ligation Master Mix was used to ligate the Stage 3 PCR products into the cloning vector pDrive according to the ligation protocol in the QIAGEN PCR Cloning Handbook 04/2001. For the transformation step, ligation mixtures were added to Subcloning Efficiency™ DH5α™ competent *E. coli* cells. The provided pUC19 (Invitrogen; diluted to 10pg/µL) was used as a positive control to verify transformation efficiency of the competent cells. The transformation reactions were then subjected to a recovery incubation (37°C, 200 rpm) with 950µL of LB broth from the Windeyer stores and finally plated on the prepared agar in either 100µL or 200µL proportion. The plates were incubated at 37°C overnight.

2.4.2. Troubleshooting

The agar plates inoculated with the transformed cells were examined for efficiency of transformation, which was poor – the blue/white colour change did not occur on many occasions – as was the cloning efficiency. Cloning was repeated another 7 times with fresh stocks of ampicillin, X-Gal, and IPTG as well as with S.O.C. medium from Invitrogen instead of LB broth. PCR was repeated with a proofreading Taq polymerase (mentioned above) and because DNA amplified this way lacks the 3' A-

overhangs necessary for effective integration into vectors for cloning, an additional Stage 4 was performed on Stage 3 PCR products following a protocol from Invitrogen (TA Cloning Kit manual, p.22) to introduce the 3' A-overhangs to the blunt DNA ends.

Although cloning of the artificial templates did not progress with satisfaction despite exhaustive troubleshooting, isolated white colonies, if any, were picked and each grown in a tube of 2mL LB broth by incubating overnight (37°C, 200 rpm). 1.5mL of the plasmid DNA from each colony was then purified with a QIAprep® Spin Miniprep Kit using the protocol for bench microcentrifuging. Following repeated sequencing failure, another purification kit and protocol was used (NucleoSpin® Plasmid, Macherey-Nagel) in anticipation that the specially-treated silica membrane would produce more purified plasmid DNA for sequencing, but this also failed.

2.4.3. Sequencing

To ensure that the DNA was free from PCR debris or contamination as well as to concentrate it for optimal sequencing results, it was cleaned using *micro*CLEAN (Microzone Ltd) following the protocol supplied. The clones of each allele were then sequenced with the BigDye® Terminator (BDT) v3.1 Cycle Sequencing Kit (Applied Biosystems) using T7 (5'-GTAATACGACTCACTATAG) and SP6 (5'-CATTTAGGTGACACTATAG) primers according to the BDTv3.1 cycle sequencing protocol. Specifically, the cycling procedures were 25 cycles of 96°C for 10s, 50°C for 5s, and 60°C for 4 minutes. The sequencing products were then purified to remove excess dye terminators following an improvised protocol (from the original BDTv3.1 guidelines) for 10μL reactions, using 125mM of ethylenediaminetetraacetic acid

(EDTA) as well as 70% and absolute ethanol. Capillary reading of the sequenced and cleaned DNA was provided by the University College London Centre for Comparative Genomics.

2.5. Endpoint genotyping with KASPar

Following the inability to achieve satisfactory results using HRM gene scanning, a KBiosciences Competitive Allele-Specific PCR SNP genotyping system (KASPar) kit was obtained (specifically the KASP indirect assay reagents, KBiosciences, Kaspar-5000). New forward (F) and reverse (R) primers, designed using KBiosciences free Primer Picker software, were ordered from Sigma Genosys and their sequences are as follow:

Minn_ALg (F1)
5'-GAAGGTCGGAGTCAACGGATTGCAAAAGTATTCAGCAGAAAGAATCAT***g***

Minn_Alt (F2)
5'-GAAGGTGACCAAGTTCATGCTAAGCAAAAGTATTCAGCAGAAAGAATCAT***t***

C1 (R1)
5'-GACCTATTACATTTTATTATTTTTTCCAGT

C2 (R2)
5'-CCTGAAGATATTTCTGTAGACCTATTACAT

The KASPar reaction mixes were made up according to instructions from the KASPar SNP Genotyping System Reagent Manual. The final Master Mix included the Assay Mix made up of the bespoke primers and the 4X Reaction Mix (see Tables 3 and 4). Specially-developed KTaq polymerase was also included in the kit. The artificial templates were from Stage 3 of the template synthesis phase and acted as positive controls, whereas the gDNA samples were drawn from C333, C341, C354, C361,

C2321, C2343, C2521, C2522, C2523, and C2548 of the control stocks in the lab. Water was included as negative control. 4µL reactions consisting of 2µL Master Mix and 2µL artificial template or gDNA (5ng/µL) were plated on to 384-well plates and subjected to the recommended PCR cycling conditions on the LightCycler® 480 Instrument (Roche):

94°C for 15m	Hot-start activation
94°C for 10s 57°C for 5s 72°C for 10s	} 20 cycles
94°C for 10s 57°C for 20s 72°C for 40s	} 18 cycles

2.5.1. Troubleshooting

Initial assays were deemed unsuccessful due to very little amplification of genotyping clusters. It was thought that the gDNA concentration was too low and/or that the Master Mix was not sufficient. Therefore, a 12.5 and 25ng/µL titration of gDNA concentration and 8µL of Master Mix was used, resulting in a total reaction volume of 10µL. It was found that there was no difference in using either concentration of gDNA, therefore all subsequent assays utilised 12.5ng/µL. The PCR cycling conditions were also amended: a third and a fourth amplification stage was added, and a read stage (where the data is captured and then displayed as cluster plots) was included after the second, third, and fourth amplification stages respectively. Reading was also done at 40°C instead of the usual 60°C. The annealing temperature was also adjusted to 56°C and then 55°C from the usual 57°C

in order to encourage amplification, but it was decided that 57°C was still optimal. At one point when acceptable genotype clustering remained elusive, 5M Betaine (in-house) and dimethyl sulfoxide (DMSO; provided in the KASPar kit) were included in the Master Mix to increase primer binding efficiency; however they did not improve clustering.

2.5.2. Successful genotyping

As an eleventh-hour effort the reverse primers designated for HRM, E1353X_TEMP_R2 and E1353X_PCR_R1, were used in place of R1 and R2 for KASPar. The optimal $MgCl_2$ concentration was 2.2mM. The results were extremely encouraging – so much so that genotyping could finally progress with using gDNA from the UCL schizophrenic cohort of 613 patients. Fresh 96-deep-well plates were made from gDNA stocks with a final concentration of 3.33ng/μL. The reaction volume was again 10μL with 4μL gDNA and 6μL Master Mix. The reactions were plated on to 384-well plates using the epMotion 5075 robot (Eppendorf) and ran using the amended protocol (4 amplification steps with annealing temperature of 57°C and 40°C reads) on the LightCycler® 480 Instrument (Roche). Duplicate plates were also examined for the 613 schizophrenic samples. Genomic DNA from 624 controls was analysed in the same way after the sequence readings for the schizophrenic samples were received, apart from the inclusion of duplicate plates as it was considered unnecessary.

2.6. Sequencing of schizophrenic and control gDNA samples

Samples (first for the schizophrenics and subsequently the controls) which did not conform to the GG homozygous cluster after endpoint genotyping analysis were selected and prepared for sequencing. Firstly the samples were amplified using optimal conditions of 2.5mM $MgCl_2$ with no Betaine (5M) and standard PCR protocol. Primers used were Exon 24 forward and reverse primers (MWG-Biotech AG), the sequences of which are:

PCM1ex22FM13R
5'-GGATAACAATTTCACACAGGATGCTTACAAAGGGTA

PCM1ex22RM13F
5'-CACGACGTTGTAAAACGACTCAAATCTAAGCTGAAA

The PCR products were then analysed using gel electrophoresis on 1% agarose gel (Sigma, agarose for routine use). HyperLadder V (Bioline) was used as the reference. As described in section 2.4.3 the DNA samples were cleaned according to the *micro*CLEAN protocol and then sequenced with the BigDye® Terminator (BDT) v3.1 Cycle Sequencing Kit as before, but using M13 promoter primers (M13F 5'-CACGACGTTGTAAAACGAC and M13R 5'-GGATAACAATTTCACACAGG) instead of SP6 and T7. The sequencing products were then purified as stated in section 2.4.3 and analysed by the Centre for Comparative Genomics, UCL.

3. Results

3.1. Artificial template synthesis

The PCR products from Stages 1 to 3 were analysed using gel electrophoresis alongside the two reference HyperLadder V shown in Figure 9 bordering the Stage1, Stage 2, and Stage 3 products, from left to right. Stage 1 products ran faster on the gel as they were smaller in size, whilst products from Stages 2 and 3 gained in size after further amplification and hence ran slower.

3.2. High resolution melting (HRM) analysis

The LightCycler® 480 Gene Scanning Software (Version 1.5) with the SYBR Green 1 detection format was used to screen for sequence variants on the Roche LightCycler® 480 Instrument. With the SensiMix HRM™ kit all wells were returned as negatives and so usage was discontinued, whilst with the new annealing temperature of 55°C three separate groups could be distinguished for the artificial templates. Unfortunately none of these assays could be considered a success in that although three distinct melting curves were obtained of the three artificial templates (Figure 10), the gDNA did not amplify or when it did, so did water (possibly due to primer-dimer formation). However the curves for the gDNA, tentatively, did conform best to the curves for the GG homozygote as was expected (Figure 11). Also, according to HRM dissociation principles, the heterozygote GT forms heteroduplices which should dissociate more readily (ie lower melting temperature) than the homozygous wild type GG and mutant TT, but this was not true of the curves obtained. Changes in the shape of the melting curves indicate the

presence of sequence variations, but the shape of the melting curve for the heterozygote did not seem readily distinguishable from that of the homozygotes.

Figure 12 shows the curves for one of the earlier assays using the recommended PCR annealing touchdown protocol. The three groups cannot be easily differentiated compared to those obtained using an annealing temperature of 55°C (Figure 10). Also, the $MgCl_2$ concentration for the former assay was 2.0mM while for the latter it was 3.0mM. This difference in quality of data illustrates that the HRM method of gene scanning is sensitive and very amenable to fine-tuning in terms of optimisation or PCR conditions.

3.3. Cloning and sequencing of artificial templates

As stated in the Methods section, despite cloning not progressing particularly well, sequencing was undertaken. The sequence was analysed using the FinchTV programme and the BLAT programme on the University of California Santa Cruz (UCSC) Genome Bioinformatics site (http://genome.ucsc.edu/). The sequences, including the templates synthesised using proofreading Taq DNA polymerase, either had extra bases apart from the expected 121 base-pairs, or were missing consecutive bases.

3.4. Endpoint genotyping with KASPar

The LightCycler® 480 Endpoint Genotyping Software (Version 1.5) with the Dual Color Hydrolysis Probe / UPL Probe detection format was used to perform data analysis on KASPar PCR products. The allele-specific dyes FAM (x-axis) and VIC (y-axis) are used to distinguish genotypes, with FAM labelling the mutant allele. A

breakthrough in this KASPar genotyping phase was the discovery of a much better separation of genotype clusters when the read was performed at 40°C instead of the recommended 60°C. The KASPar probes each contain a hairpin loop structure which prevents unincorporated probes from fluorescing. Analysing the data at 60°C results in increased background fluorescence (discussed later). Figure 13 is included to drive the point home.

Endpoint genotyping analyses performed on the 613 schizophrenic DNA samples yielded the cluster plots seen in Figures 14 and 15. Two schizophrenic DNA samples were found to contain the 4057G→T mutation, namely S0155 and S2363, and these were used in subsequent assays as positive controls. In Figure 12 (top), although two of the negative controls (ie water) appeared to have amplified as much as 1/3 of the DNA samples, thus emitting a fluorescence signal of about 4.0, the DNA sample labelled S0155 had clearly separated from the main cluster smear. Results from the duplicate plate (Figure 14, bottom), although yielding less satisfactory clusters, confirmed that S0155 was indeed a definite candidate to be sequenced. Figure 15 depicts cluster plots for duplicates of the second batch of schizophrenic DNA samples – on these plates the artificial templates GG, GT, and TT were included as references, and it was apparent that the sample S2363 clustered closer to the GT template. The group calling function provided by the Endpoint Genotyping Software did not prove to be extremely reliable in these cases.

Further endpoint genotyping with KASPar was conducted with the 624 control DNAs. The schizophrenic samples found to be positive were included on the plates. As shown in Figure 16, the bulk of the control samples were called as one

cluster, while the schizophrenic samples S0155 and S2363 were called differently although one of the S2363 samples from the first batch was called as a negative.

3.5. Sequencing of schizophrenic and control gDNA samples

Sequencing reads were analysed using the Staden Package Pregap4 software following the protocol delineated in Table 5. S0081 was used as the forward and reverse reference for sequence analysis. Sequencing outliers from the KASPar assays for schizophrenics returned 2 positive samples which, as has been mentioned above, were S0155 and S2363. Mutations are tagged by the mutation scanning software in the "Contig Editor" shown in Figure 17, and the trace displays for S0155 and S2363 are shown in Figure 18. Apart from them, the other samples which separated from the main cluster during endpoint genotyping analysis turned out only to be false positives. However the sequences for 8 samples were returned as negative reads and were therefore included to be re-sequenced with the controls, as were S0155 and S2363. As of the time of writing, sequencing of these samples had not been completed.

4. Discussion

The main aim of this research project was to replicate a finding of a single nucleotide polymorphism or SNP mutation named the 4057G→T mutation (Sawa et al., 2008, Personal communication) on exon 24 of the PCM1 gene, leading to increased susceptibility to schizophrenia. This mutation was originally found by Sawa and colleagues in 1 out of 32 cases and the mutation segregated with schizophrenia-spectrum disorder in the proband's family. In the UCL sample of 613 schizophrenics studied for this project, 2 cases were found to possess the mutation of interest.

4.1. The inability to capitalise on High Resolution Melting analysis

The original intention was to use High Resolution Melting (HRM) to characterise base pair changes in PCR amplicons by studying melting curves. This, in effect, scans all the bases for SNP mutations before confirmation by sequencing. The HRM gene scanning analysis is a good method for detecting genetic variation. It was chosen for this project because it does not rely on synthesising expensive fluorescently-labelled primers. Artificial templates corresponding to the DNA sequence around the 4057G→T mutation were designed in order to ensure reliable detection of the mutation by HRM analysis. However sequences obtained after cloning and sequencing of the templates indicated that the desired templates had not been created. These incorrect sequences could have resulted from problems encountered during cloning or sequencing, where failure remains unexplained. It is possible that the template concentration was very low, as discovered via nanodrop quantification, and therefore this affected cloning and subsequent sequencing.

Time constraints led to testing of the HRM method on artificial templates whose sequence had not been confirmed by DNA sequencing. Templates representing both the GG and TT homozygotes and the GT heterozygote revealed three distinct melt curves at an annealing temperature of 55°C. It was unfortunate that the heteroduplex GT did not melt earlier than the homoduplex TT and especially the wild type GG as per HRM principles. Problems encountered with amplifying the genomic DNA (gDNA) samples were insurmountable. It is uncertain as to why they did not amplify as well as observed for the artificial templates. Even when amplification was achieved there was evidence of contamination because negative "water only" controls also amplified. The gDNA and $MgCl_2$ concentrations were titrated, PCR cycling conditions amended, and HRM was repeated with a different chemistry (SensiMix HRM™ EvaGreen®), but the gDNA samples did not yield satisfactory melting curves. Having said that however, the gDNA curves did fall nearest to the curves for the GG templates, and perhaps with more time the HRM assay could have been optimised for both the artificial templates and the gDNA samples. The use of different primer pairs could also improve HRM results.

4.2. Reverting back to endpoint genotyping

Since the melting profile for the artificial templates were clearly different, it was decided that the regions with the SNP were sufficiently well-characterised for endpoint genotyping analysis. Initial assays using the KASPar chemistry and specially-designed oligonucleotides, one for each allele of the SNP, did not yield acceptable cluster plots even with all troubleshooting steps taken, for example the addition of the PCR enhancing agents Betaine and DMSO. Both agents encourage

strand separation (Frackman, Kobs, Simpson, & Storts, 1998). Characteristic of the complex chemistry involved, the reverse primers intended for HRM did however work in place of the ones designed for endpoint genotyping. These sets of primers were designed to anneal to different regions around the 4057G→T mutation.

Performing a read at 40°C instead of the usual 60°C greatly improved the clustering of different genotypes. This advance was possible due to a thorough understanding of the basic mechanisms of KASPar endpoint genotyping – at higher temperatures the unincorporated fluorescently-labelled primers begin to lose their looped structure, and therefore the fluorescence level is higher because the fluorescence of the reporter dye is no longer suppressed by the quencher dye. The high level of background fluorescence might have been detrimental to the current assays and thus led to poor cluster calling at 60°C.

4.3. The finding of the 4057G→T mutation and its implications

2 out of 613 schizophrenic DNA samples were discovered to carry the 4057G→T mutation after sequencing. The reason why only 2 cases were positive out of the entire cohort as compared to 1 out of 32 cases in the original study is that those 32 cases had been reported to have excess allele sharing among affected family members over the chromosome 8p22 linked region. This mutation, although very rare, seems so far to be highly penetrant and exerts a dominant effect compared to the normal allele on the other chromosome. Homozygous mutant TT genotypes would be very rare and unlikely to be observed in the sample sizes studied so far. There is also the possibility that foetuses with the TT genotype do not survive.

The G to T base change introduces a premature stop codon, named E1353X (Sawa et al., 2008), which would lead to either truncation or nonsense-mediated decay of the protein. The substituting T allele, in other words, would result in a mutant protein with a gain in abnormal function or loss of function of the normal protein. As the PCM1 gene encodes centrosomal proteins and is essential for cell division by regulating microtubular dynamics, loss of its usual function could lead to neurodevelopmental disorders. It is likely that this mutation on the PCM1 gene has serious consequences because it disrupts not only the normal functions of this protein but leads to further effects in the neuronal migration pathway by virtue of its interaction with such proteins as DISC1, Huntington Associated Protein (HAP1), Bardet Biedel Syndrome (BBS) proteins, Ninein, and Pericentrin (Datta et al., submitted). DISC1 has multiple functions at the centrosome, mitochondria, postsynaptic densities, and the nucleus (Ichizuka et al., 2006), while it seems that PCM1 has a more restricted centrosome-associated function. HAP1 is a cytoskeletal protein bound to PCM1 and has a role in vesicular trafficking (Engelender et al., 1997; Li et al., 1996).

Of particular note is that one (S0155) of our 2 cases, for whom the family pedigree (Figure 1) is known, has a mother suffering from a schizophrenia related disorder. All siblings also had been treated at some point with mental health problems. This is reminiscent of the finding that this heterozygous 4057G→T mutation segregated with schizophrenia-spectrum disorders in the proband's family in the original study. Bearing in mind Sawa et al.'s (2008) suggestion that mutations on PCM1 may cause a subtype of schizophrenia, namely deficit schizophrenia characterised primarily by negative symptoms, it is notable that relatives of deficit

probands are thought to be at a higher risk of developing subclinical features resembling deficit schizophrenia. They also may be more prone to developing schizophrenia. However this remains contentious and further studies are warranted (Kirkpatrick et al., 2000).

Another interesting point for discussion is the brain morphology study by Gurling and colleagues (2006), which found significant grey matter deficits in the orbitofrontal cortex in cases who had an inherited PCM1 susceptibility to schizophrenia. This region is preferentially indicated in reward-related processing and motivational behaviours and abnormalities here could bias symptomatology towards affective and behavioural features. Supporting research (Kendler et al., 2000) reports the finding that families with linkage to 8p21-22 had significantly more affective deterioration, more thought disorder, and overall a poorer outcome compared to non-8p21-22-linked families. However, as Kendler et al. have acknowledged, their conclusions are at best tentative due to important methodological limitations. In addition they found fewer depressive symptoms in the 8p21-22-linked families (Kendler et al., 2000), which is one of the characteristics of patients with deficit schizophrenia (Kirkpatrick et al., 2001). Deficit schizophrenics also have a lower risk of affective disorders (Kirkpatrick et al., 2000). Whether schizophrenics with mutations on PCM1 are more likely to suffer from the deficit form of the disease remains to be seen, but the notion of them being specific to a subtype of schizophrenia seems promising.

The ages-at-onset of disease in the current 2 cases are 17 and 18 respectively. Approximately 12-33% of schizophrenics develop their illness onset before 18 years of age (Krausz & Muller-Thomsen, 1993; Loranger, 1984) and they

would be classified as suffering from early-onset schizophrenia-spectrum disorders (Kumra et al., 2008). When compared with adult-onset schizophrenic patients, those with early-onset disorders suffer, long before first manifestation of psychotic symptoms, from more premorbid abnormalities in language and motor functions as well as some symptoms similar pervasive developmental disorder (Hollis, 1995). Within the current context, this can be reconciled with the role PCM1 plays in the maturing brain with regard to development of the nervous system and neuronal activity.

Early-onset disorder patients also appear to have higher rates of psychiatric co-morbidities like major depressive disorder and attention-deficit hyperactive disorder (Ross et al., 2006). Relevant to this is that the 8p21-22 chromosome locus has been linked to mood disorders apart from schizophrenia (Lewis et al., 2003; Chiu et al., 2002; Blouin et al., 1998). Accordingly, it could be that this particular mutation on PCM1 predisposes to early-onset disorders, arguably a more severe manifestation of schizophrenia (Rapoport & Inoff-Germain, 2000), as well as affective disorders such as bipolar disorder and depression. This provides genetic data in support of a common link among clinically distinct disorders which nonetheless affect similar neurobiological processes (Cantor & Geschwind, 2008), which can be attributed to the importance of PCM1 in ensuring centrosome integrity and thus normal neuronal migration.

A pathophysiological model simulating reduced synaptic connectivity due to loss of neuritic processes put forward by McGlashan and Hoffman (2000) predicts that younger age of onset would lead to more debilitating deterioration post-onset and that hallucinations and other positive symptoms would be less persistent than

negative and disorganised symptoms. This seems an attractive model insofar as the current mutation and gene function are concerned.

One last issue to be discussed is the apparent rarity of this mutation in schizophrenics. Recently Walsh and associates (2008), using copy number variation (CNV) analysis, showed that multiple and individually rare mutations alter genes in neurodevelopmental pathways and contribute to schizophrenia. Indeed, it has been suggested that heterogeneous rare mutations play an important role in the genetics of schizophrenia (McClellan et al., 2006), and it is proposed that this PCM1 mutation is just one of them.

Conclusions and future directions

In summary, the finding of a heterozygous mutation on exon 24 of the PCM1 gene has been replicated in at least 2 cases out of 613. It is proposed that this mutation, although rare, has a high disease penetrance and could be associated with more severe young-onset schizophrenia and schizophrenia-spectrum disorders with primarily negative symptoms. It will be of interest to see at what age the proband in the original study developed schizophrenia and if future cases carrying this mutation also are patients with early-onset schizophrenia-spectrum disorders. In-depth investigations into the clinical features of the probands and their relatives should also be conducted.

Genotyping and sequencing of the control DNA samples were only conducted after the discovery of these 2 positive samples in order to save costs. Due to the lack of time, sequencing of the controls had not been completed by the end of the current project. Therefore, before progressing with further studies, it

should be confirmed that this mutation is not present in any of the controls. The next step forward will then be to perform family linkage analysis on the family of the 2 cases to see if the mutation segregates with schizophrenia. Possible future attempts include introducing this 4057G→T mutation in rats – to create a model for schizophrenia – and to investigate the effects on the developing brain.

The current finding of a rare mutation causing susceptibility to schizophrenia will need further replication. Such efforts should be vigorously pursued because of its importance to the whole of schizophrenia research. There is an urgent need to find better treatment and preventative measures against schizophrenia. As such, a recent finding of interest is that the antipsychotic clozapine down-regulated PCM1 gene expression while haloperidol did not (Rizig et al., submitted). This could be consistent with the proposal that PCM1 genetic mutations cause a distinct subtype of schizophrenia. The current PCM1 finding may be the first concrete example of a high penetrance mutation associated with schizophrenia.

8,856 words (including Tables and Figures / Diagrams)

References

Ansley S.J., Badano J.L., Blacque O.E., Hill J., Hoskins B.E., Leitch C.C, Kim, J.C., Ross, A.J., Eichers, E.R., Teslovich, T.M., Mah, A.K., Johnsen, R.C., Cavender, J.C., Lewis, R.A., Leroux, M.R., Beales, P.L., & Katsanis, N. (2003). Basal body dysfunction is a likely cause of pleiotropic Bardet-Biedl syndrome. *Nature, 425*(6958), 628-633.

Bengra, C., Mifflin, T.E., Khripin, Y., Manunta, P., Williams, S.M., Jose, P.A., & Felder, R.A. (2002). Genotyping of essential hypertension single-nucleotide polymorphisms by a homogeneous PCR method with universal energy transfer primers. *Clinical Chemistry, 48*(12), 2131-2140.

Blouin, J.L., Dombroski, B.A., Nath, S.K., Lasseter, V.K., Wolyniec, P.S., Nestadt, G., Thornquist, M., Ullrich, G., McGrath, J., Kasch, L., Lamacz, M., Thomas, M.G., Gehrig, C. Radhakrishna, U., Snyder, S.E., Balk, K.G., Neufeld, K., Swartz, K.L., DeMarchi, N., Papadimitriou, G.N., Dikeos, D.G., Stefanis, C.N., Chakaravarti, A., Childs, B., Housman, D.E., Kazazian, H.H., Antonarakis, S., Pulver, A.E. (1998). Schizophrenia susceptibility loci on chromosomes 13q32 and 8p21. *Nature Genetics, 20*(1), 70-73.

Brzustowicz, L.M., Honer, W.G., Chow, E.W., Little, D., Hogan, J., Hodgkinson, K., *et al.* (1999). Linkage of familial schizophrenia to chromosome 13q32. *American Journal of Human Genetics, 65*(4), 1096-1103.

Cantor, R.M. & Geschwind, D.H. (2008). Genome, interrupted. *Neuron, 58*, 165-167.

Carpenter, W.T. Jr., Heinrichs, D.W., & Wagman, A.M. (1988). Deficit and nondeficit forms of schizophrenia: the concept. *American Journal of Psychiatry, 145*, 578-583.

Chiu, Y.F., McGrath, J.A., Thornquist, M.H., Wolyniec, P.S., Nestadt, G., Swartz, K.L., Lasseter, V.K., Liang, K.Y., & Pulver A.E. (2002) Genetic heterogeneity in schizophrenia II: conditional analyses of affected schizophrenia sibling pairs provide evidence for an interaction between markers on chromosome 8p and 14q, *Molecular Psychiatry, 7*(6), 658-664.

Datta, S.R., McQuillin, A., Rizig, M., Blaveri, E., Thirumalai, S., Kalsi, G., Lawrence, J., Bass, N.J., Puri, V., Choudhury, K., Pimm, J., Crombie, C., Fraser, G., Walker, N., Curtis, D., Zvelebil, M., Pereira, A., Kandaswamy, R., St. Clair, D., Gurling, H.M.D. A threonine to isoleucine mis-sense mutation in the pericentriolar material 1 (PCM1) gene is strongly associated with schizophrenia; *submitted.*

Engelender, S., Sharp, A.H., Colomer, V., Tokito, M.K., Lanahan, A., Worley, P., , E.L.F., & Ross, C.A. (1997). Huntingtin-associated protein 1 (HAP1) interacts with the p150Glued subunit of dynactin. *Human Molecular Genetics, 6*(13), 2205-2212.

Frackman, S., Kobs, G., Simpson, D., & Storts, D. (1998). Betaine and DMSO: Enhancing agents for PCR. *Promega Notes, 65*, 27.

Gurling, H.M., Kalsi, G., Brynjolfson, J., Sigmundsson, T., Sherrington, R., Mankoo, B.S., *et al.* (2001). Genomewide genetic linkage analysis confirms the presence of susceptibility loci for

schizophrenia, on chromosomes 1q32.2, 5q33.2, and 8p21-22 and provides support for linkage to schizophrenia, on chromosomes 11q23.3-24 and 20q12.1-11.23. *American Journal of Human Genetics, 68*(3), 661-673.

Gurling, H.M., Critchley, H., Datta, S.R., McQuillin, A., Blaveri, E., Thirumalai, S., Pimm, J., Krasucki, R., Kalsi, G., Quested, D., Lawrence, J., Bass, N., Choudhury, K., Puri,V., O'Daly, O., Curtis, D., Blackwood, D., Muir, W., Malhotra, A.K., Buchanan, R.W., Good, C.D., Frackowiak, R.S., Dolan, R.J. (2006). Genetic association and brain morphology studies and the chromosome 8p22 pericentriolar material 1 (PCM1) gene in susceptibility to schizophrenia. *Archives of General Psychiatry, 63* (8), 844-854.

Harrison, P.J. & Weinberger, D.R. (2005). Schizophrenia genes, gene expression, and neuropathology: on the matter of their convergence. *Molecular Psychiatry, 10*, 40–68.

Hollis, C. (1995). Child and adolescent (juvenile onset) schizophrenia. A case control study of premorbid developmental impairments. *British Journal of Psychiatry, 166*, 489–495.

Ichizuka, K., Paek, M., Kamiya, A., Sawa, A. (2006). A review of Disrupted-In-Schizophrenia-1 (DISC1): neurodevelopment, cognition, and mental conditions. *Biological Psychiatry, 59* (12), 1189-1197.

Kamiya, A., Kubo, K., Tomoda, T., Takaki, M., Youn, R., Ozeki, Y., Sawamura, N., Park, U., Kudo, C., Okawa, M., Ross, C.A., Hatten, M.E., Nakajima, K., & Sawa, A. (2005). A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. *Nature Cell Biology, 7*(12), 1167-1178.

Kendler, K.S., MacLean, C.J., O'Neill, F.A., Burke, J., Murphy, B., Duke, F., *et al.* (1996). Evidence for a schizophrenia vulnerability locus on chromosome 8p in the Irish Study of High-Density Schizophrenia Families. *American Journal of Psychiatry, 153*(12), 1534-1540.

Kendler, K.S., Myers, J.M., O'Neill, F.A., Martin, R., Murphy, B., MacLean, C.J., Walsh, D., & Straub, R.E. (2000). Clinical features of schizophrenia and linkage to chromosomes 5q, 6p, 8p, and 10p in the Irish Study of High-Density Schizophrenia Families. *American Journal of Psychiatry, 157*, 402-408.

Kim, J.C., Badano, J.L., Sibold, S., Esmail, M.A., Hill, J., Hoskins, B.E., Leitch, C.C., Venner, K., Ansley, S.J., Ross, A.J., Leroux, M.R., Katsanis, N., & Beales, P.L. (2004). The Bardet-Biedl protein BBS4 targets cargo to the pericentriolar region and is required for microtubule anchoring and cell cycle progression. *Nature Genetics, 36*(5), 462-470.

Kirkpatrick, B., Ross, D.E., Walsh, D., Karkowski, L., & Kendler, K.S. (2000). Family characteristics of deficit and nondeficit schizophrenia in the Roscommon Family Study. *Schizophrenia Research, 45*, 57-64.

Kirkpatrick, B., Buchanan, R.W., Ross, D.E., & Carpenter, W.T. Jr. (2001). A separate disease within the syndrome of schizophrenia. *Archives of General Psychiatry, 58*, 165-171.

Kubo, A., Sasaki, H., Yuba-Kubo, A., Tsukita, S., & Shiina, N. (1999). Centriolar satellites: molecular characterization, ATP-dependent movement toward centrioles and possible involvement in ciliogenesis. *Journal of Cell Biology, 147*(5), 969-980.

Kubo, A., Tsukita, S. (2003). Non-membranous granular organelle consisting of PCM-1: subcellular distribution and cell-cycle-dependent assembly/disassembly. *Journal of Cell Science, 116*(Pt 5), 919-928.

Krausz, M. & Muller-Thomsen, T. (1993). Schizophrenia with onset in adolescence: an 11-year follow-up. *Schizophrenia Bulletin, 19*, 831–841.

Kumra, S., Oberstar, J.V., Sikich, L., Findling, R.L., McClellan, J.M., Vinogradov, S., & Schulz, S.C. (2008).Effects and tolerability of second-generation antipsychotics in children and adolescents with schizophrenia. *Schizophrenia Bulletin, 34* (1), 60-71.

Lewis, C.M., Levinson, D.F., Wise, L.H., DeLisi, L.E., Straub, R.E., Hovatta, I., Williams, N.M., Schwab, S.G., Pulver, A.E., Faraone, S.E., Brzustowicz, L.M., Kaufmann, C.A., Garver, D.L., Gurling, H.M., Lindholm, E., Coon, H., Moises, H.W., Byerley, W. Shaw, S.H., Mesen, A., Sherrington, R., O'Neill, F.A., Walsh, D., Kendler, K.S., Ekelund, J., Paunio, T., Lonnqvist, J., Peltonen, L., O'Donovan, M.C., Owen, M.J., Wildenauer, D.B., Maier, W., Nestadt, G., Blouin, J.L., Antonarakis, S.E., Mowry, B.J., Silverman, J.M., Crowe, R.R., Cloninger, C.R., Tsuang, M.T., Malaspina, D., Harkavy-Friedman, J.M., Svrakic, D.M., Bassett, A.S., Holcomb, J., Kalsi, G., McQuillin, A., Brynjolfson, J., Sigmundsson, T., Petursson, H., Jarzin, E., Zoega, T., & Helgason, T. (2003). Genome scan meta-analysis of schizophrenia and bipolar disorder, part II: Schizophrenia. *American Journal of Human Genetics, 73* (1), 34-48.

Li, X.-J., Sharp, A.H., Li, S.-H., Dawson, T.M., Snyder, S.H., & Ross, C.A. (1996). Huntingtin-associated protein (HAP1): Discrete neuronal localizations in the brain resemble those of neuronal nitric oxide synthase. *Proceedings of the National Academy of Sciences, 93*(10), 4839-4844.

Loranger, A. (1984). Sex difference in age at onset of schizophrenia. *Archives of General Psychiatry, 41*, 157–161.

McClellan, J.M., Susser, E., King, M.C. (2006). Maternal famine, de novo mutations, and schizophrenia.*The Journal of the American Medical Association, 296*, 582-584.

McGlashan, T.H. & Hoffman, R.E. (2008). Schizophrenia as a disorder of developmentally reduced synaptic connectivity. *Archives of General Psychiatry, 57*, 637-648.

Miyoshi, K., Asanuma, M., Miyazaki, I., Diaz-Corrales, F.J., Katayama, T., Tohyama, M., & Ogawa, N. (2004). DISC1 localises to the centrosome by binding to kendrin. *Biochemical and Biophysical Research Communications, 317*(4), 1195-1199.

Online Mendelian Inheritance in Man, OMIM (TM). Johns Hopkins University, Baltimore, MD. MIM Number: {600299}: {6/18/2002}: World Wide Web URL: http://www.ncbi.nlm.nih.gov/omim/

Rapoport, J.L. & Inoff-Germain, G. (2000). Update on Childhood-onset Schizophrenia. *Current Psychiatry Reports, 2*, 410-415.

Rizig, M.A., McQuillin, A., Robinson, M., Hunt, S.P., Ng, A., Harrison, A., *et al.* (2009). A gene expression and systems pathway analysis of the effects of clozapine compared to haloperidol in the mouse brain implicates susceptibility genes for schizophrenia *Pharmacogenomics; submitted.*

Ross, R., Heinlein, S., & Tregellas, H. (2006). High rates of comorbidity are found in childhood-onset schizophrenia. *Schizophrenia Research, 88*, 90–95.

Saha, S., Chant, D., Welham, J., McGrath, J.A. (2005). A systematic review of the prevalence of schizophrenia. *PLoS Medicine, 2*, 413-433.

Spitzer, R. & Endicott, J. (1977). *The schedule for affective disorders and schizophrenia, lifetime version*. 3rd ed. New York State Psychiatric Institute: New York.

Suarez, B.K., Duan, J., Sanders, A.R., Hinrichs, A.L., Jin, C.H., Hou, C., Buccola, N.G., Hale. N., Weilbaecher, A.N., Nertney, D.A., Olincy, A., Green, S., Schaffer, A.W., Smith, C.J., Hannah, D.E., Rice, J.P., Cox, N.J., Martinez, M., Mowry, B.J., Amin, F., Silverman, J.M., Black, D.W., Byerley, W.F., Crowe, R.R., Freedman, R., Cloninger, C.R., Levinson, D.F., & Gejman, P.V. (2006). Genomewide linkage scan of 409 European-ancestry and African American families with schizophrenia: Suggestive evidence of linkage at 8p23.3-p21.2 and 11p13.1-q14.1 in the combined sample. *American Journal of Human Genetics, 78* (2), 315-333.

Sullivan, P.F. (2008). Schizophrenia genetics: the search for a hard lead. *Current Opinion in Psychiatry, 21*, 157-160.

Walsh, T., McClellan, J.M., McCarthy, S.E., Addingtong, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A., Stray, S.M., Rippey, C.F., Roccanova, P., Makarov, V., Lakshmi, B., Findling, R.L., Sikich, L., Stromberg, T., Merriman, B., Gogtay, N., Butler, P., Eckstrand, K., Noory, L., Gochman, P., Long, R., Chen, Z., Davis, S., Baker, C., Eichler, E.E., Meltzer, P.S., Nelson, S.F., Singleton, A.B., Lee, M.K., Rapoport, J.L., King, M.-C., Sebat, J. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science, 320*, 539.

Tables

In Methods, Section 2.2:

Table 1. Sequences of primers used to synthesis artificial templates.

GG homozygous template	TT homozygous template
Common forward and reverse primers	
E1353X_TEMP_F1 5'-TCAGCCCAGACTGAAGAGCCTGTTCAAGCAAAAGTATTCAGCAGAAAG	
E1353X_PCR_F1 5'-GAACAGACATTCAGCCCAGA	
E1353X_TEMP_R2 5'-CTGAAGATATTTCTGTAGACCTATTACATTTTATTATTTTTTCCAGTTGC	
E1353X_PCR_R1 5'-CATACCTGAAGATATTTCTGTAGACCT	
G allele-specific primer	T allele-specific primer
E1353X_TEMPG_R1 5'-CCAGTTGCT*c*ATGATTCTTTCTGCTG	E1353X_TEMPT_R1 5'-CCAGTTGCT*a*ATGATTCTTTCTGCTG

Table 2. Components used in preparation of the artificial template over 3 stages.

Component	Volume (μL)		
	Stage 1	Stage 2	Stage 3
10x *Pfx* Amplification Buffer	5	5	5
10mM dNTP mixture	6	6	6
50mM MgSO	1	1	1
E1353X_TEMP_FI	1.5	n/a	n/a

E1353X_TEMPG/T_R1	1.5	n/a	n/a
E1353X_PCR_F1	n/a	1.5	1.5
E1353X_TEMP_R2	n/a	1.5	n/a
E1353X_ PCR_R1	n/a	n/a	1.5
Template DNA (1/100 of product from previous stage)	n/a	1.5	1.5
Platinum® *Pfx* DNA Polymerase	0.4	0.4	0.4
PCR water (Sigma®)	34.6	33.1	33.1
Total volume	**50.0**	**50.0**	**50.0**

In Methods, Section 2.5:

Table 3. Components of the Assay Mix.

Assay Mix(1/10)	Volume in µl
F1 primer	1.2
F2 primer	1.2
R1 or R2 primer	3
H_2O	94.6
Total volume	**100**

Table 4. Components of the Master Mix excluding Assay Mix.

	A	B	C
384-well	**2.2mM mgcl2**	**1.8mM mgcl2**	**2.5mM mgcl2**
DNA	2	2	2
4X Reaction Mix	1	1	1
Assay Mix 1/10	0.55	0.55	0.55
KTaq polymerase 1/10	0.13	0.13	0.13
$MgCl_2$	0.032	0	0.056
H_2O	0.288	0.32	0.264
Total volume	**4**	**4**	**4**

In Results, Section 3.5:

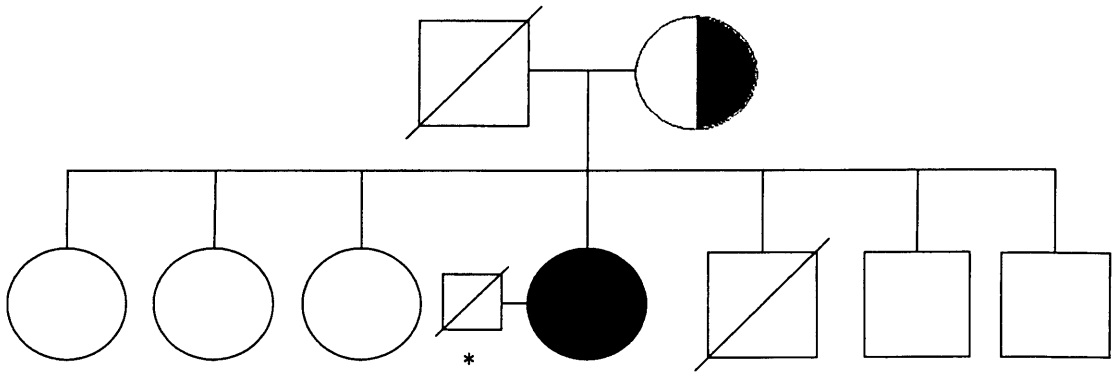
Table 5. Protocol used for sequencing analysis with Staden Package Pregap4.

Sequencing analysis with Staden Package Pregap4
1. Download sequencing reads from designated Wolfson Institute website.
2. Scan through sequences using programmes like FinchTV or Trev 1.9 and choose a good quality forward and reverse reference. Rename as ref_primername_F or ref_primername_R.
3. Refer to the University of California Santa Cruz (UCSC) Genome Bioinformatics site (http://genome.ucsc.edu/) for reference sequence of exon of interest by feeding in sequence of forward and reverse primers used. Save as a text file called primernoreference into same folder as data files.
4. Load Pregap4 and add files to be processed by selecting all files after changing the file type to "all*.*". Files added should include reference text file from UCSC.
5. Configure modules by turning on all options except "Phred" and "Interactive clipping".
6. Adjust settings for: Reference traces and sequences • Reference Strand +ve = file copied and named reference_primername_F • Reference Strand -ve = file copied and named reference_primername_R • Reference sequence = text file from UCSC and saved as primernoreference Gap4 shotgun Assembly Create new database and name it.
7. Run.
8. Remove bad quality sequences by performing left and right quality control.
9. Open Database (aux file) in Gap4 and analyse for mutations. "Tags" mark bases where the mutation scanning software thinks there are mutations.
10. Double click on "Tags" to open the "Trace display" window, which will show the test sample trace, the reference trace, and a difference trace.

Figures and Diagrams

In Discussion, Section 4.3:

Figure 1. Pedigree diagram for S0155, sourced from clinical files. The genotype of the proband is shown – wt/E1353X – where E1353X denotes a premature termination codon.



Key:

Normal male

Normal female

Deceased male

Female with manic depression or schizophrenia

Female with other types of depression or schizoid/ paranoid/ schizotypal personality

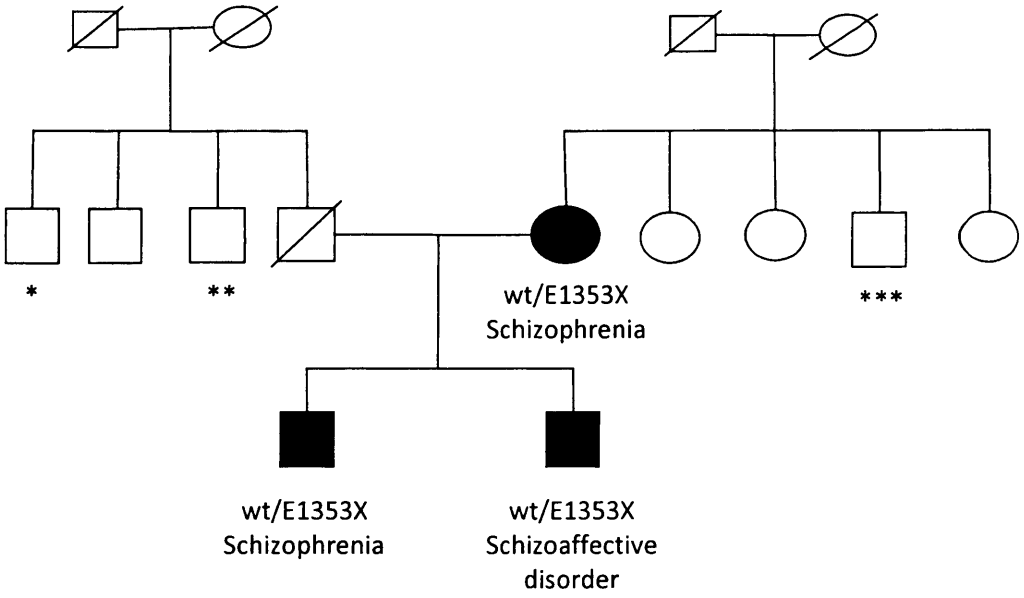
Other comments: All other siblings have been treated at some time for mental health problems.

*Husband of proband, death by suicide

**Death by suicide, described as having "blunted" affect

In Introduction:

Figure 2. Family pedigree showing individuals with the heterozygous E1353X mutation (genotype wt/E1353X). Asterisks denote members with a psychiatric phenotype. Adapted from Sawa et al. (2008).



Key:

Normal male

Male with manic depression or schizophrenia

Normal female

Female with manic depression or schizophrenia

Deceased male

Deceased female

Other comments:

*Major depression

**Alcohol abuse

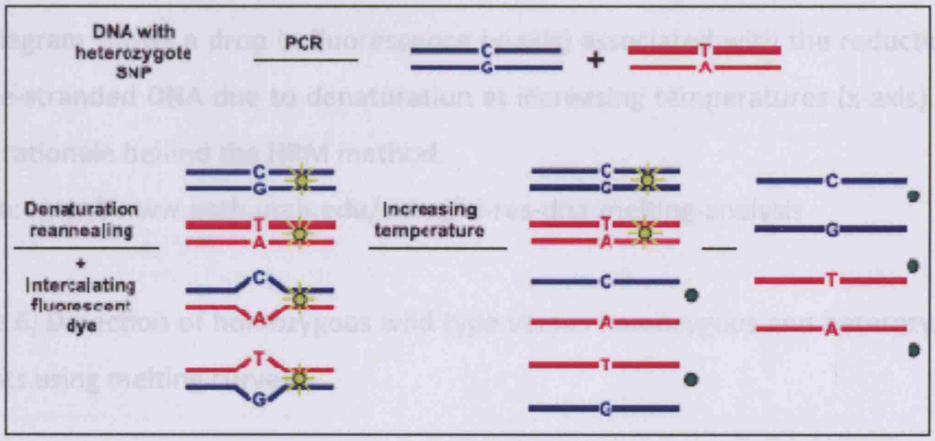
*** Substance abuse

Figure 3. Pericentriolar-material 1 (PCM1) with red line marking the 8p21-22 locus.

Sourced from the UCSC Genome Browser on Human March 2006 Assembly using

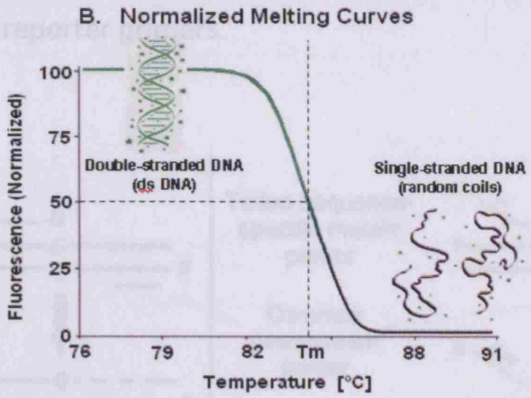
Section 1.1.1:

Figure 4. Diagram from LightCycler® 480 Gene Scanning Software (Version 1.5) manual depicting the rationale behind the high resolution melting (HRM) analysis.



DNA containing the SNP of interest undergoes PCR and is then available in high copy number. PCR is performed in the presence of an intercalating and saturating dye which binds only to double-stranded DNA (dsDNA). Therefore, at the beginning of a melting experiment, the initial fluorescence is high, and at increasing temperature the dsDNA dissociates into single-stranded DNA (ssDNA) and fluorescence diminishes. Bases which are not complementary (ie C-A or T-G) form weaker bonds and therefore dissociate earlier.

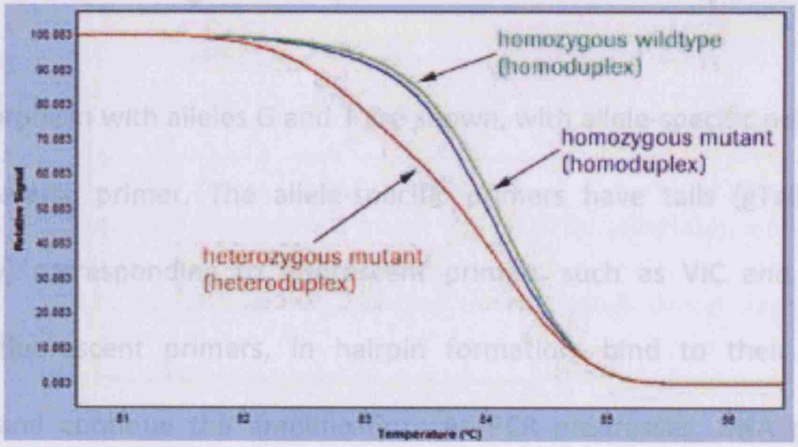
Figure 5. A normalised melting curve of dissociating double-stranded DNA



The diagram shows a drop in fluorescence (y-axis) associated with the reduction in double-stranded DNA due to denaturation at increasing temperatures (x-axis). This is the rationale behind the HRM method.

Source: http://www.path.utah.edu/news/hi-res-dna-melting-analysis

Figure 6. Detection of homozygous wild type versus homozygous and heterozygous variants using melting curves.

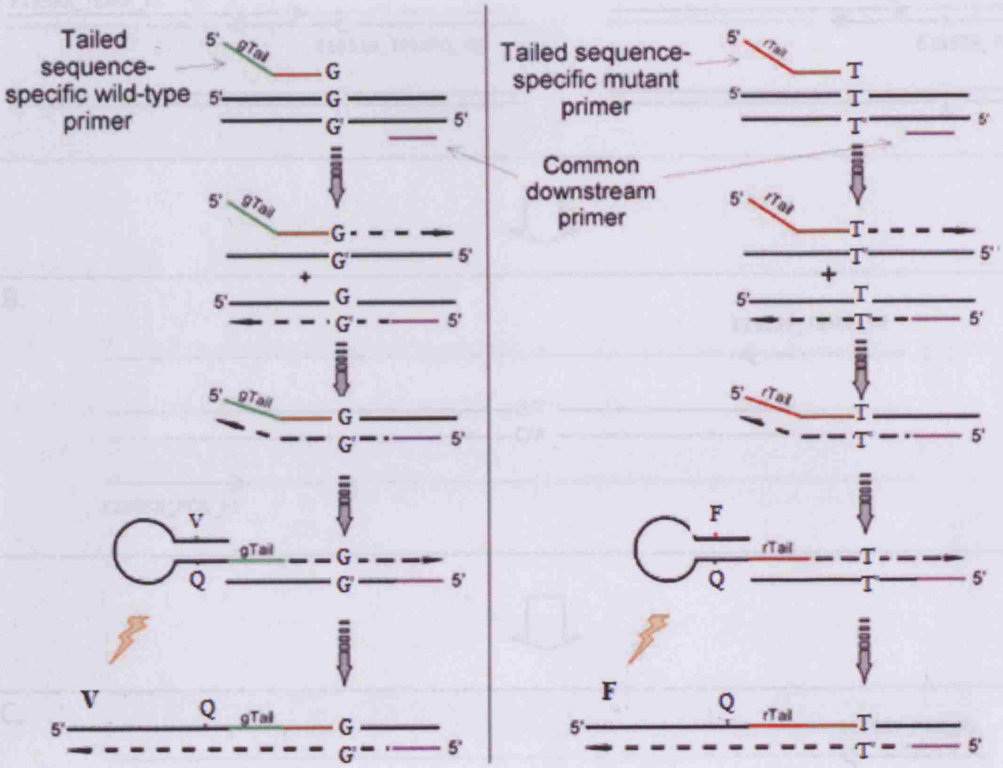


Heterozygous mutants form heteroduplices which dissociate at a lower melting temperature and have melting curves which obviously differ from that of the wild type. HRM is sensitive enough to pick up differences between wild type and mutant homozygotes. The y-axis shows the relative fluorescence signal and the x-axis shows the temperature.

Source: LightCycler® 480 Gene Scanning Software (Version 1.5) manual

Section 1.1.2:

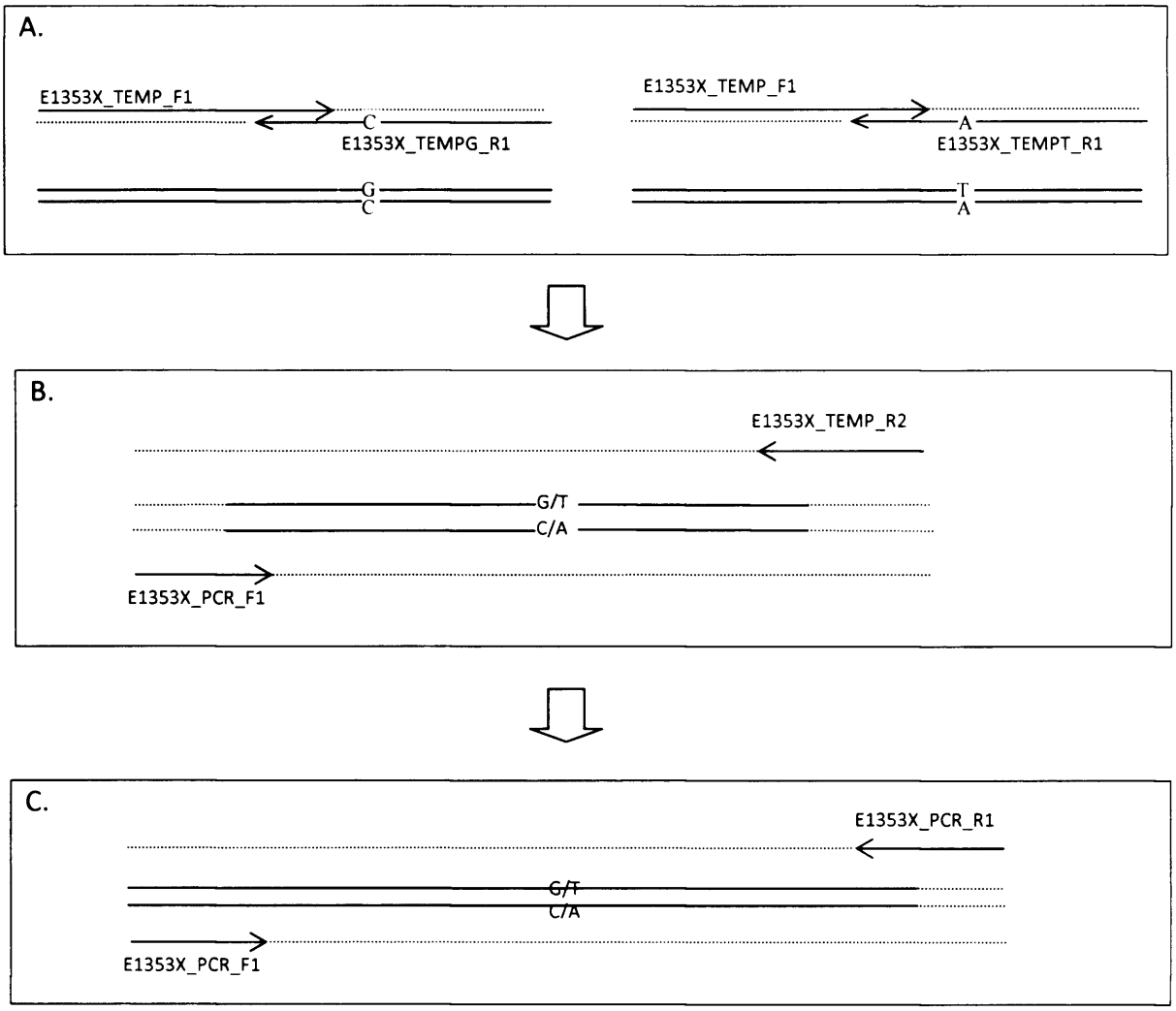
Figure 7. Overview of the SNP detection mechanism with allele-specific primers and universal fluorescent reporter primers.



The polymorphism with alleles G and T are shown, with allele-specific primers and a common reverse primer. The allele-specific primers have tails (gTail and rTail respectively) corresponding to fluorescent primers such as VIC and FAM. The quenched fluorescent primers, in hairpin formation, bind to their respective amplicons and continue the amplification. As PCR progresses, DNA polymerase opens up the hairpin structures and separates the reporter molecules (F , FAM, and V, VIC) from the quencher (Q). This allows the reporters to emit fluorescence when excited. Dashed lines and arrows show the directions of DNA polymerase extension. Modified from Figure 1, Bengra et al. (2002).

In Methods:

Figure 8. Preparation of artificial templates for High Resolution Melting (HRM).



Boxes A to C illustrate the 3 stages via which the artificial templates were prepared. The primers used in each stage are included in the figure. Box A depicts Stage 1 whereby the allele-specific primers overlap and were used to create the GG and TT homozygotes respectively. In the subsequent stages products from the previous stage served as templates and common primers, which overlapped with parts of the templates, were used. Arrows and bold lines represent the primers, while dotted lines show extension of the PCR products. Modified from figure courtesy of Dr Andy McQuillin.

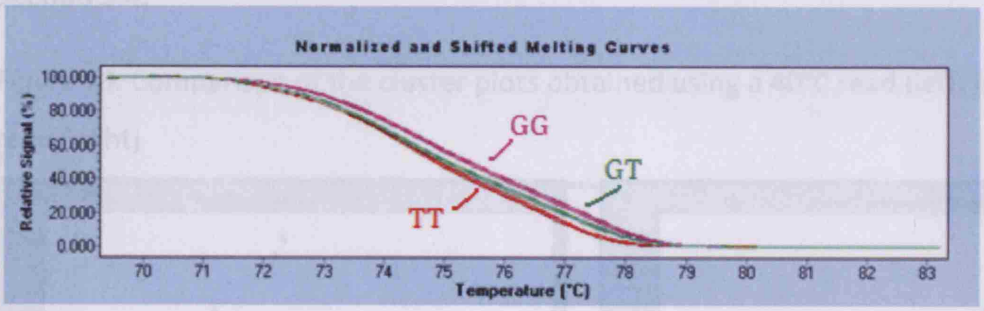
In Results:

Section 3.1:

Figure 9. Gel electrophoresis of PCR products from Stages 1 to 3 with HyperLadder V reference on extreme left and right.

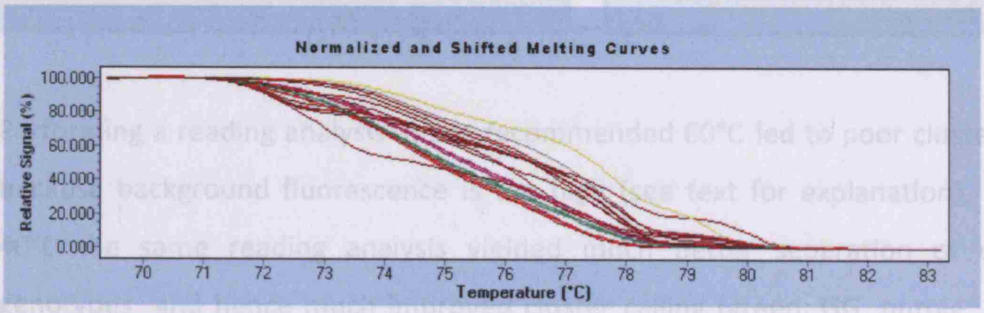
Section 3.2:

Figure 10. Normalised and temperature-shifted melting curves for the 3 templates obtained using an annealing temperature of 55°C.



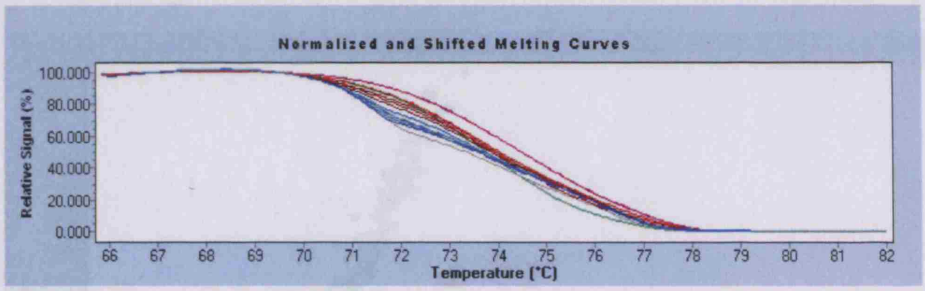
Each genotype differs by a single base, resulting in a slightly different melting temperature and thus generating distinguishable melting curves (pink: GG, green: GT, red: TT).

Figure 11. Normalised and temperature-shifted melting curves for the 3 templates and genomic DNA (gDNA) samples.



Identical to the graph above but also showing melting curves for gDNA samples. Though not ideal, they do lie closest to the curve for the wild type GG (pink).

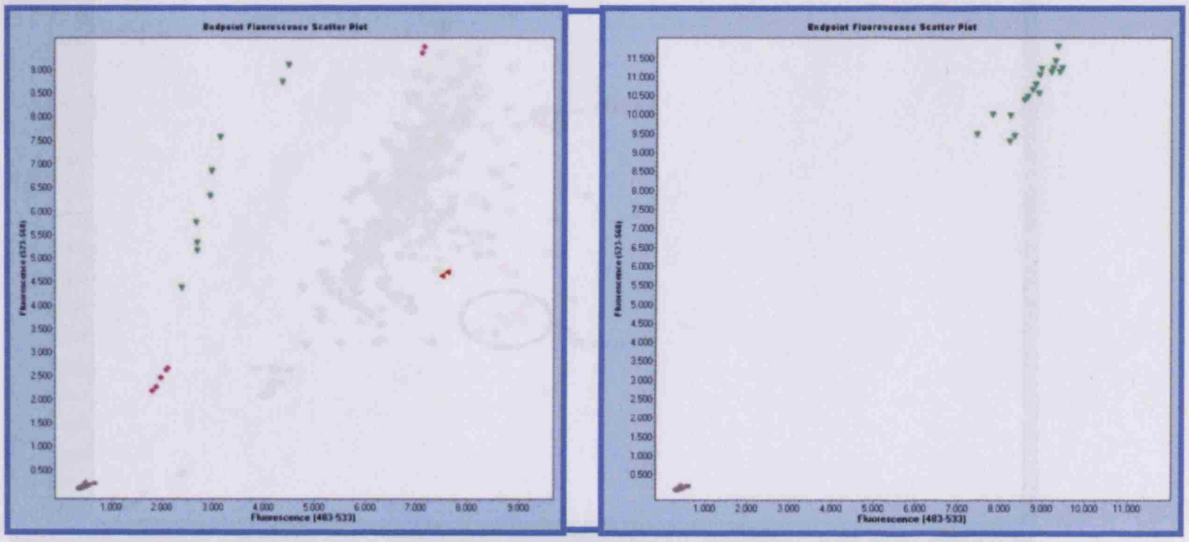
Figure 12. Normalised and temperature-shifted melting curves for the 3 artificial templates obtained using the recommended PCR annealing touchdown protocol.



Melting curves shown here were obtained using an annealing temperature of 65°C to 53°C. The curves are noticeably less tight compared to those in Figure 10 and the 3 different genotypes were not called as 3 separate groups by the software.

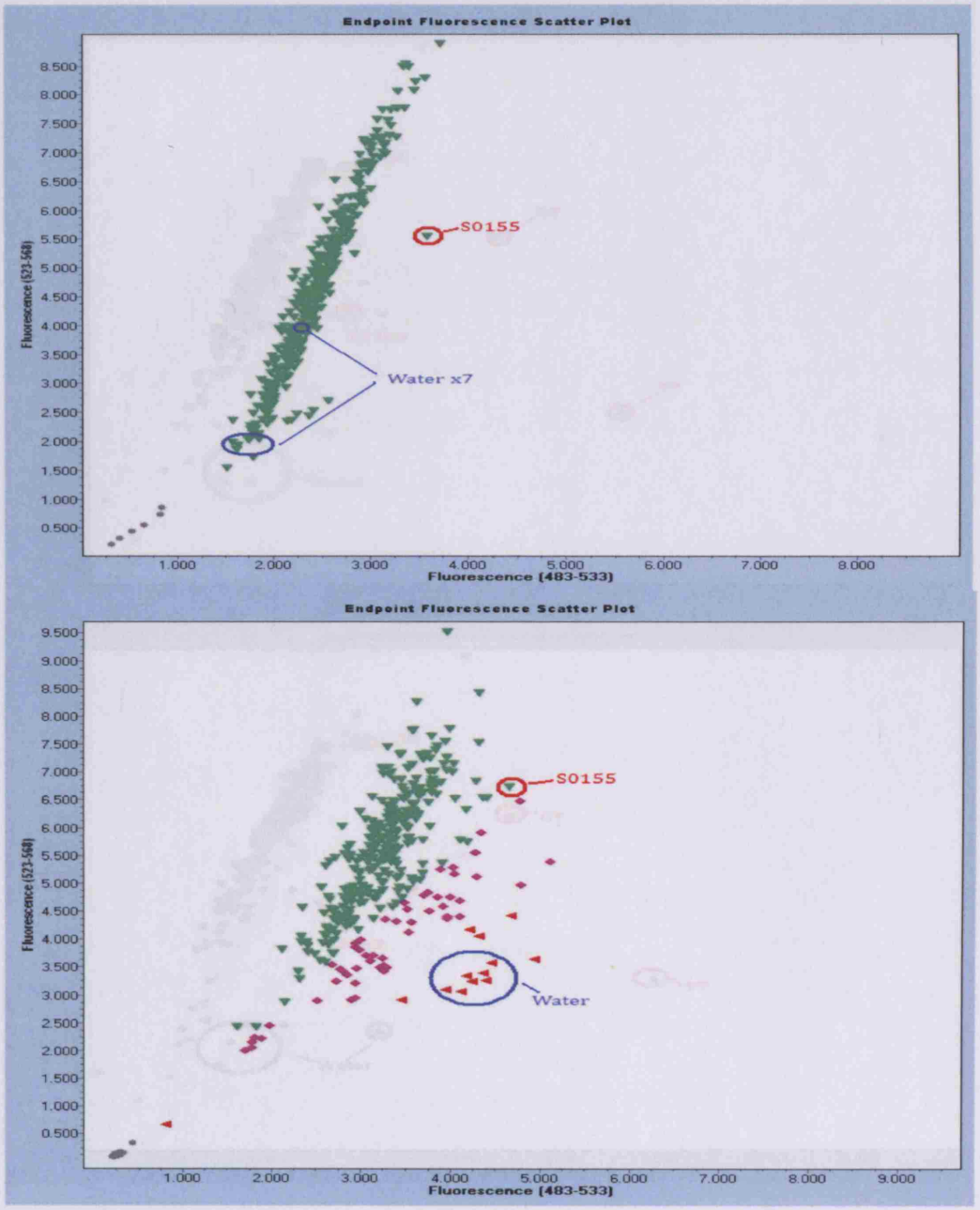
Section 3.4:

Figure 13. Comparison of the cluster plots obtained using a 40°C read (left) and 60°C read (right)



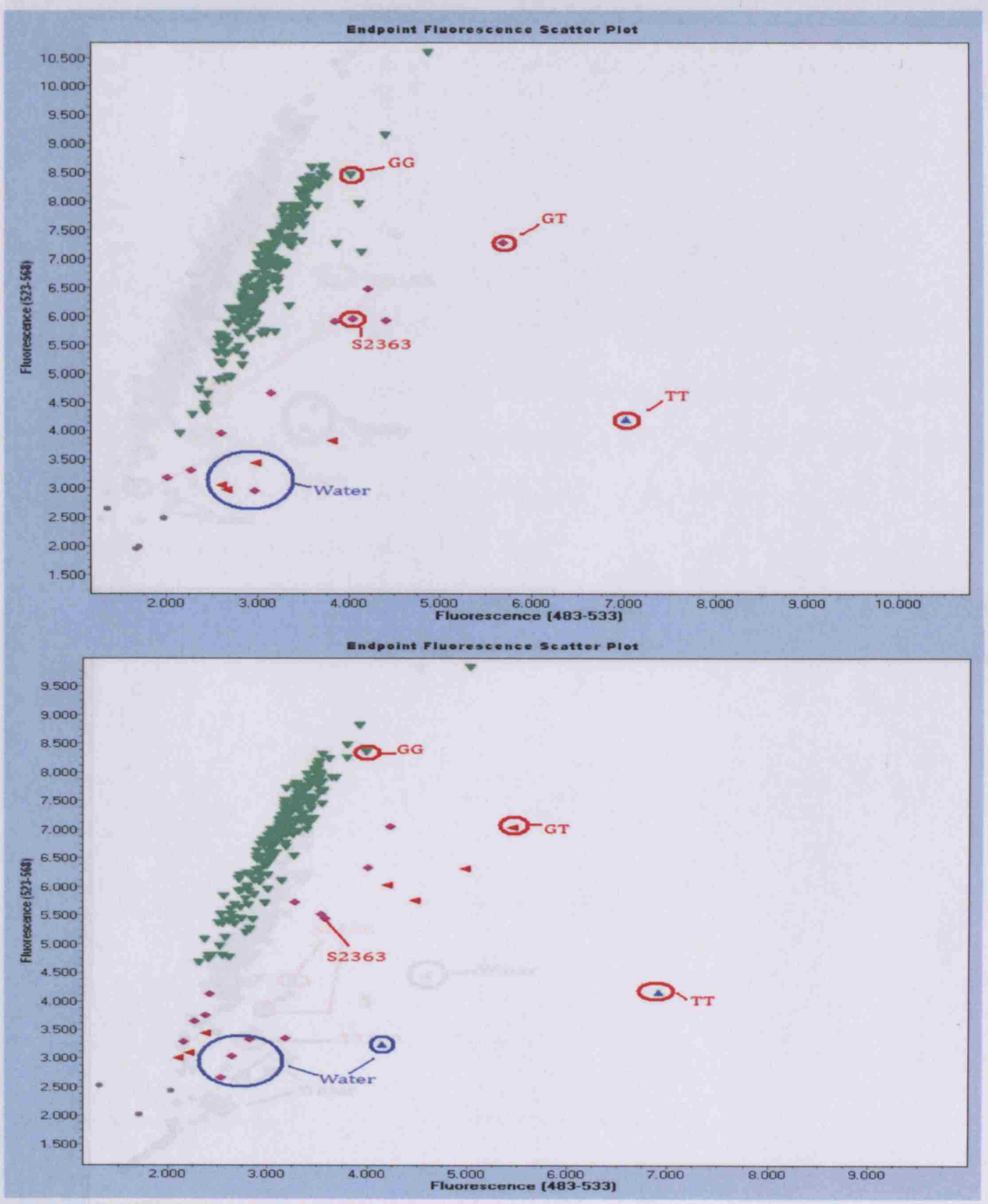
Performing a reading analysis at the recommended 60°C led to poor cluster calling because background fluorescence is too high (see text for explanation), while at 40°C the same reading analysis yielded much better separation of different genotypes and hence much improved cluster calling (green: GG, purple: GT, red: TT). Axes show the fluorescence.

Figure 14. Endpoint fluorescent cluster plots for the first batch of 613 schizophrenic DNA samples.



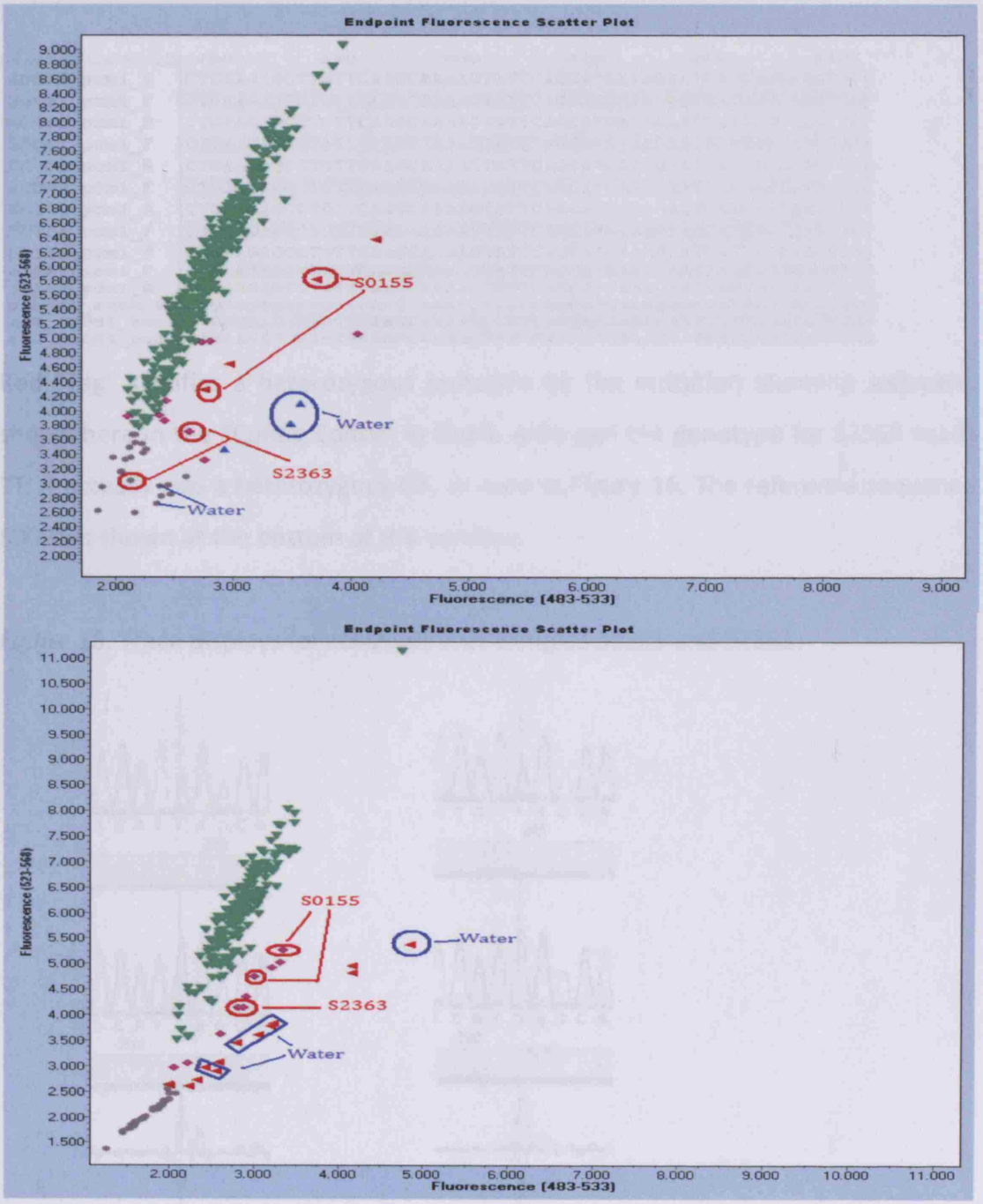
These cluster plots were obtained from duplicate plates. In both, the schizophrenia DNA sample S0155, circled in red, has moved away from the main GG (green) cluster. 7 negative controls (water) are circled in blue. The grey dots in the bottom left corner of the graphs are samples which did not amplify.

Figure 15. Endpoint genotyping cluster plots for the second batch of 613 schizophrenic DNA samples.



On both plates, the bottom one a duplicate, the 3 artificial templates were included as reference and there were 4 negative controls each. The schizophrenic DNA sample S2363 clustered near the GT template in both instances.

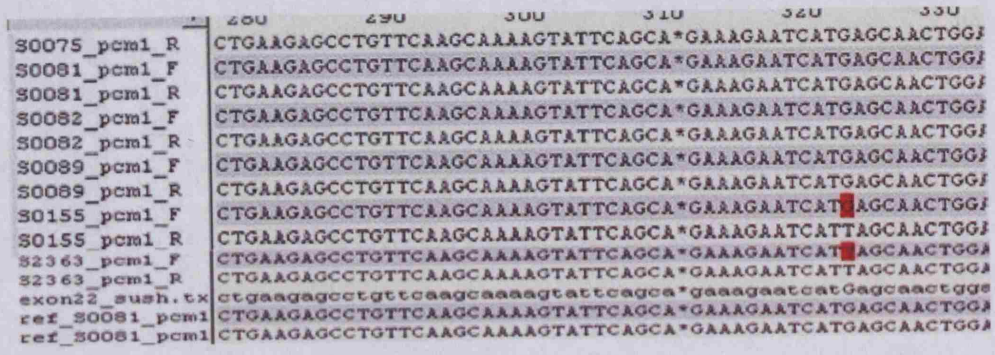
Figure 16. Endpoint fluorescence cluster plot for the first (top) and second (bottom) batches of 624 control DNA samples.



Control DNA samples shown here with the majority in a tight green cluster. The schizophrenic DNA samples S0155 and S2363 (circled in red), found to contain the 4057G→T mutation after sequencing, were also included in the assay as positive controls. Water was the negative control (circled in blue).

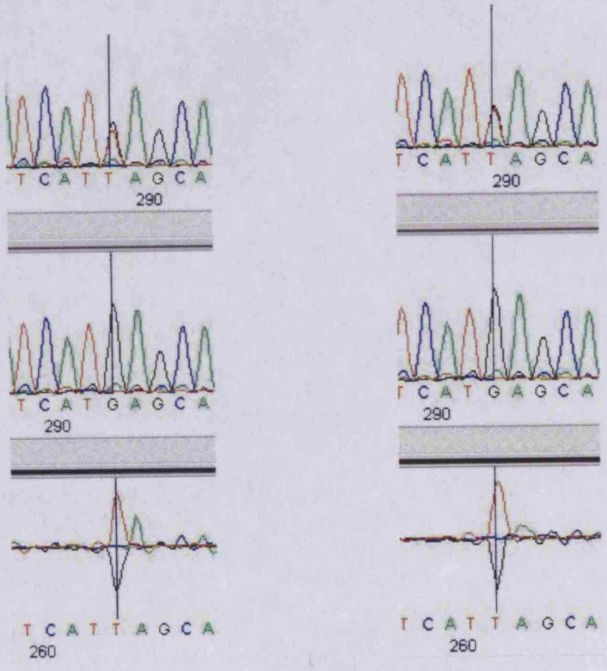
Section 3.5:

Figure 17. "Contig Editor" window in Gap4.



Red "Tag" signifies a heterozygous mutation by the mutation scanning software, shown here in the "Contig Editor" in Gap4. Although the genotype for S2363 reads TT, it actually was a heterozygous GT, as seen in Figure 16. The reference sequence S0081 is shown at the bottom of the window.

Figure 18. Trace displays for schizophrenia samples S0155 and S2363.



As illustrated in the trace displays, both S0155 (left) and S2363 (right) contain the heterozygous 4057G→T mutation. From top to bottom, the traces shown are the test sample trace, the reference trace, and the difference trace.