Disruption of two novel genes by a translocation co-segregating with schizophrenia

Citation for published version:

Digital Object Identifier (DOI):
10.1093/hmg/9.9.1415

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Human Molecular Genetics

Publisher Rights Statement:
This is a publisher-authenticated version of the article. It is also available online at:
http://hmg.oxfordjournals.org/content/9/9/1415

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Disruption of two novel genes by a translocation co-segregating with schizophrenia

J. Kirsty Millar¹⁺, Julie C. Wilson-Annan¹⁸, Susan Anderson¹, Sheila Christie¹, Martin S. Taylor¹, Colin A. M. Semple¹, Rebecca S. Devon¹, David M. St Clair², Walter J. Muir¹,³, Douglas H. R. Blackwood¹,³ and David J. Porteous¹

¹Medical Genetics Section, Department of Medical Sciences, The University of Edinburgh, Molecular Medicine Centre and MRC Human Genetics Unit, both at Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK,
²The Department of Mental Health, University Medical Building, The University of Aberdeen, Foresterhill, Aberdeen AB9 2ZD, UK and ³Department of Psychiatry, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF, UK

A balanced (1;11)(q42.1;q14.3) translocation segregates with schizophrenia and related psychiatric disorders in a large Scottish family (maximum LOD = 6.0). We hypothesize that the translocation is the causative event and that it directly disrupts gene function. We previously reported a dearth of genes in the breakpoint region of chromosome 11 and it is therefore unlikely that the expression of any genes on this chromosome has been affected by the translocation. By contrast, the corresponding region on chromosome 1 is gene dense and, not one, but two novel genes are directly disrupted by the translocation. These genes have been provisionally named Disrupted-In-Schizophrenia 1 and 2 (DISC1 and DISC2). DISC1 encodes a large protein with no significant sequence homology to other known proteins. It is predicted to consist of a globular N-terminal domain(s) and helical C-terminal domain which has the potential to form a coiled-coil by interaction with another, as yet, unidentified protein(s). Similar structures are thought to be present in a variety of unrelated proteins that are known to function in the nervous system. The putative structure of the protein encoded by DISC1 is therefore compatible with a role in the nervous system. DISC2 apparently specifies a non-coding RNA molecule that is antisense to DISC1, an arrangement that has been observed at other loci where it is thought that the antisense RNA is involved in regulating expression of the sense gene. Altogether, these observations indicate that DISC1 and DISC2 should be considered formal candidate genes for susceptibility to psychiatric illness.

INTRODUCTION

Schizophrenia is a serious and debilitating disease affecting ~1% of the population worldwide. There is compelling evidence from family, twin and adoption studies for a significant genetic basis to the disease (1). This has initiated searches directed at identification of the genetic component using methods such as linkage analysis, association studies of candidate genes and mapping of cytogenetic abnormalities in psychiatric patients, procedures which have been applied successfully to monogenic disorders. Psychiatric illnesses are more complex, however, and apparently result from the combined effects of multiple genes, with inheritance complicated by environmental factors (1). Consequently no genes involved in the aetiology of such illnesses have yet been definitively identified.

We are studying a large Scottish family (2) in which a balanced translocation segregates with major mental illness (maximum LOD = 6.0; D.H.R. Blackwood, A. Fordyce, M. Walker, E. Drysdale, J.K. Millar, D.M. St Clair, D.J. Porteus and W.J. Muir, manuscript in preparation), based on the hypothesis that the rearrangement has directly disrupted gene function, leading to psychosis. This family may be atypical due to the wide spectrum of disorders present (schizophrenia, schizoaffective disorder, bipolar affective disorder, unipolar affective disorder and adolescent conduct disorder). However, it is likely that identification of the genetic factor(s) involved in the aetiology of disease in these patients would facilitate discovery and understanding of the underlying genetic defects in unrelated psychotic individuals.

RESULTS

To clone the chromosome 1 breakpoint, a 2.5 kb EcoRI fragment from chromosome 11, containing the site of the breakpoint (3), was used to screen an EcoRI genomic library constructed from a translocation cell line. A 2.7 kb EcoRI fragment, corresponding to the translocation fragment from the derived 1 chromosome was obtained (Fig. 1A). This fragment was used to rescreen the same library, yielding a 7.3 kb clone, containing the site of the chromosome 1 breakpoint (Fig. 1A). The derived 11 fragment was obtained using the polymerase
chain reaction (PCR). Alignment of breakpoint sequence from all four fragments (chromosome 11, chromosome 1, derived 11, derived 1, GenBank accession nos AF222984, AF222983, AF222986 and AF222985, respectively) (Fig. 1B) shows that the translocation resulted in replacement of TCAG with AA on the derived 11 chromosome, and that, consequently, no major rearrangement of sequence accompanied the translocation event.

Sequence analysis provided no evidence for the presence of any transcripts within the sequence surrounding the chromosome 11 breakpoint. However, the chromosome 1 breakpoint fragment sequence includes three expressed sequence tags (ESTs: AA249072, W04811 and D78808) and an exon from a messenger RNA, AB007926 (Fig. 2).

We have provisionally named the anonymous mRNA AB007926 DISC1 (Disrupted-In-Schizophrenia 1), because analysis of the chromosome 1 breakpoint sequence indicates that the gene is directly disrupted by the translocation, which takes place within an intron (Fig. 2). Furthermore, this analysis indicates that the direction of transcription is proximal to distal.

Using a combination of cDNA library screening and RACE (rapid amplification of cDNA ends), we obtained 6913 nt of cDNA sequence transcribed from the DISC1 gene (accession no. AF222980), which matched AB007926 (99.985%) over 6833 nt, and extended 14 nt further 5′. Four different 5′ RACE products, our longest cDNA clone and AB007926 all terminate within 14 nt, suggesting that this is the true 5′ end of the transcript. Furthermore, Nucleotide Identity X (NIX, http://menu.hgmp.mrc.ac.uk/
Menu-bin/Nix/Nix.pl) analysis of 2.4 kb of genomic sequence encompassing the 5' end of DISC1 identified a 758 nt putative CpG island (70% GC) containing 120 nt from the 5' end of the transcript, and two potential promoters 55 and 359 nt upstream. The first ATG in DISC1 is located at position 54 and is not part of a strong translation initiation consensus (4). The predicted coding sequence, starting at the first ATG, consists of 2565 nt, with a stop codon at position 2616. The 3' untranslated region (UTR) is 4294 nt extending to a poly(A) tail. A consensus polyadenylation signal is located at position 6892, 16 nt upstream of the poly(A) tail. The gene is tagged at the 3' end by UniGene cluster Hs.26985.

Sixty-six nucleotides (2295–2360) are not contained within the putative coding sequence of AB007926. This deletion corresponds to a common alternative splicing event (Fig. 3, Table 1). DISC1 is present as a major transcript of ~7.5 kb (Fig. 4A and B) in all adult tissues examined. The size discrepancy between the sequence we have obtained and the transcripts detectable on northern blots may be due to polyadenylation. DISC1 was not detected on northern blots of fetal tissues (data not shown), although reverse transcription (RT)-PCR experiments indicate that fetal transcripts do exist (Table 1).

The open reading frame (ORF) in DISC1 encodes a putative protein of 854 amino acids. Protein structure prediction programmes (http://dodo.cpmc.columbia.edu/predictprotein/) suggest that DISC1 can be divided into two distinct regions of secondary structure (Fig. 5A). The N-terminal region (amino acids 1–347) is predicted to consist of one or more globular domains. The C-terminal region is predicted to consist entirely of α-helix interspersed with several short loops, and contains regions with the potential to form coiled-coils, structures that arise when the helical stretches of separate proteins interact. Alternative splicing introduces an additional loop into the C-terminal domain.

BLAST 2.0 and FASTA 3 searches of the SWall database (SwissProt plus TREMBL) at the European Bioinformatics Institute (http://www2.ebi.ac.uk/) reveal matches (~21% identity, 41% similarity across 500 amino acids) between the α-helical region of DISC1, particularly around the stretches of coiled-coil, and the known or predicted coiled-coil domains of...
several other proteins. These similarities are likely to result from biased sequence composition imposed by coiled-coil structure and therefore probably reflect structural resemblances between the proteins.

**DISC2** (accession no. AF222981) was identified from the ESTs at the chromosome 1 breakpoint (Fig. 2) and extended by cDNA library screening, RACE and RT–PCR. This identified 15 178 nt of contiguous genomic sequence which is known to be transcribed. The transcript 3′ end is tagged by Unigene cluster Hs.96883. There are two consensus polyadenylation signals at positions 15 072 and 15 161, 107 and 18 nt upstream of the poly(A) tail.

The 5′ end of the transcript has not yet been located, but **DISC2** so far consists of a single large exon encompassing the translocation breakpoint and the 189 nt exon of **DISC1**. Sequence analysis indicates that the direction of transcription is distal to proximal. **DISC2** therefore overlaps with **DISC1**, but is transcribed in the opposite orientation. **DISC2** transcripts are most abundant in heart where species of >9.5 kb, and of ~6, 3 and 2.5 kb are present (Fig. 4C). RT–PCR indicates that **DISC2** is also transcribed in several fetal tissues (Table 1) although transcripts were not detectable on northern blots (data not shown).

No significant ORF has been identified within 15 178 nt of the **DISC2** transcript. The longest ORF deduced to be present encodes 57 amino acids, while the start codon of this ORF is not in a good context (4), suggesting that the sequence lacks any protein coding potential. Furthermore, a survey of the EMBL database (ftp://ftp.ebi.ac.uk/pub/databases/embl/release/) indicates that the longest known 3′ UTR in a human gene is 9280 nt (doublecortin), while the average length is 2131.83 nt ± 1368.09 (based on 861 full-length 3′ UTR sequences). Therefore, if **DISC2** is a protein-coding gene, it possesses a 3′ UTR in excess of 15 kb, substantially outside the size range of other known 3′ UTRs. In addition, **DISC2** possesses certain similarities to the 17 572 nt non-coding mRNA-like transcript **NTT** (5), as summarized (Fig. 5B). These observations suggest that **DISC2** is a non-coding structural RNA gene, although final confirmation awaits identification of the transcript 5′ end.

---

**Table 1.** RT–PCR analysis of **DISC1** and **DISC2**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Age (weeks)</th>
<th><strong>DISC1</strong></th>
<th><strong>DISC2</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>proximal</td>
<td>distal</td>
<td>proximal</td>
</tr>
<tr>
<td>Brain</td>
<td>8.3</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td></td>
<td>10.3</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td></td>
<td>13.3</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td>Heart</td>
<td>8.8</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td></td>
<td>9.1</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td></td>
<td>9.3</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td>Kidney</td>
<td>10.0</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td>Spleen</td>
<td>14.8</td>
<td>+</td>
<td>+ (2)</td>
</tr>
<tr>
<td>Limb</td>
<td>10.3</td>
<td>+</td>
<td>+ (2)</td>
</tr>
</tbody>
</table>

Approximate ages of gestation are given in weeks. 2, two bands obtained using one primer pair indicating alternative splicing; +, transcript detected; –, transcript not detected.
We have cloned and sequenced the breakpoints of a (1;11)(q42.1;q14.3) translocation linked to schizophrenia and related psychosis (2), and identified two novel genes, DISC1 and DISC2, both disrupted by the translocation. One or both of the disrupted alleles may be responsible for the psychiatric disorders suffered by carriers of the translocation. Furthermore, these genes may also be involved in the mental illness of patients unrelated to the family segregating the translocation. However, no independent evidence for a locus in this region of the genome has yet been presented, although recent reports of suggestive linkage to 1q32 and 1q32–41 in patients suffering from bipolar disorder and schizophrenia, respectively (6,7), are intriguing.

DISC1 protein matches are essentially restricted to structural similarities to myosins, structural proteins and proteins that are involved in motility and/or transport (particularly microtubule binding proteins). In the context of psychiatric illness, it is interesting to note that many of these proteins (dynactin, D-CLIP-190, citron, post-synaptic density proteins, FEZ1 and hyaluronan receptor, for example) are implicated in processes such as axon guidance, synaptogenesis, functioning of the synapse and intracellular transport along axons and dendrites (8–15). It is an intriguing possibility that the function of DISC1 is similar, suggesting a role in development of the nervous system and/or neuronal activity, and therefore adding further to the evidence pointing towards involvement in the aetiology of mental illness.
By analogy to many other examples of mammalian genes with endogenous antisense RNA transcripts (16–18), DISC2 presents an attractive mechanism by which DISC1 expression may be regulated. There is evidence that antisense RNAs affect expression of the sense gene (i) in the nucleus at the levels of transcription, RNA processing or export from the nucleus, or (ii) in the cytoplasm by influencing RNA stability or translation. Similarly, DISC2 might act at one of these levels to influence expression of DISC1.

In translocation carriers, transcription of both DISC1 and DISC2 is predicted to occur from their endogenous promoters on the derived chromosomes, unless unidentified transcription signals of major effect are removed or inactivated by the rearrangement. This contention is supported by preliminary experiments which detect transcription of DISC1 from the derived 1 chromosome (data not shown). It is unlikely that corresponding truncated transcripts lacking a 5′ end would be produced as well, because the breakpoint region of chromosome 11 is apparently transcriptionally inactive (3,19).

Several scenarios resulting from production of 3′ truncated DISC1 and DISC2 transcripts could be envisaged, but three are most obvious. First, truncated DISC1 protein may be produced. It would lack the C-terminal 257 amino acids, including one of the regions of most strongly predicted coiled-coil forming potential. This would be expected to reduce the overall coiled-coil forming potential and stability as a multimer, while retaining the unidentified function of the globular N-terminus. Production of a partially active protein could conceivably result in a dominant-negative effect. Second, truncated DISC2 transcripts would retain complementarity to the normal full-length DISC1 transcripts. However, DISC2 truncation could affect events following interaction between truncated DISC2 and normal DISC1 transcripts. Therefore, if DISC2 does regulate expression of DISC1 it is possible that this mechanism would be negatively affected, resulting in dysregulation of both the non-translocated and translocated alleles, and a pleiotropy of dominant-negative effects. Third, irrespective of the putative regulatory role of DISC2 on DISC1, the gene product may have independent (regulatory) functions affected by truncation.

We propose that alteration of DISC1 and/or DISC2 activity, by truncation and/or by abnormal regulation of expression, is causally linked to the psychiatric illness in translocation carriers. Dysregulation of the novel functions of DISC1 and DISC2 in the absence of a translocation event may play a more general role in susceptibility to psychiatric illness.

MATERIALS AND METHODS

Cell culture

The lymphoblastoid cell line MAFLI from an individual bearing the t(1;11)(q21.1;q14.3) translocation, somatic cell hybrids MIST7.4 and MIST39 bearing the derived 1 or derived 11 translocation chromosomes respectively, and their culture conditions, have been described previously (20). On the derived chromosome 1, DNA has been lost from 1q42.1-qter and replaced with chromosome 11 material from 11q14-3qter. The derived 11 chromosome is the reciprocal translocated chromosome. The cell line A9(NEO-1)-4, a mouse A9 hybrid cell line carrying human chromosome 1, and its culture requirements, have been reported previously (21).

DNA preparation

Human and λ genomic DNA was prepared by standard methods (22). Plasmid and cosmID DNA was prepared using Qiagen (Crawley, UK) plasmid midi kits.

RNA extraction and cDNA synthesis

Human fetal tissues were obtained from the United Kingdom Medical Research Council Tissue Bank. Total RNA was extracted using RNazol B (Biogenesis, Poole, UK) according to the manufacturer’s instructions. First strand cDNA synthesis was carried out on DNase I-treated RNA using the random hexamer primer from the SUPERSCRIPT Preamplification System (Gibco BRL, Paisley, UK) according to the manufacturer’s instructions. One microlitre of the resulting cDNA was used in standard PCRs.

Genomic library construction and screening

Genomic DNA from the translocation cell line MAFLI was digested with EcoRI, ligated into EcoRI-digested and dephosphorylated AZAPII (Stratagene, Cambridge, UK), and packaged using Gigapack Gold II packaging extract (Stratagene) according to the manufacturer’s instructions. Bacteriophage were plated using Escherichia coli XL1-Blue MRF+ and the library of clones screened using standard methods. Excision of clones from the λ vector was carried out as advised by the manufacturer, releasing genomic fragments cloned into pBluescript SK(–). The library was screened using a 2.15 kb repeat-free HindIII–EcoRI sub-fragment of the 2.5 kb EcoRI fragment containing the chromosome 11 breakpoint (3), followed by the 2.7 kb derived 1 fragment.

cDNA library screening

5′-STRETCH PLUS cDNA libraries of 20–26 week fetal brain and 20–25 week fetal heart, constructed in bacteriophage λgt10 and gt11, respectively, were obtained from Clontech (Basingstoke, UK) and screened according to the manufacturer’s instructions. Inserts were obtained from pure clones using two methods. First, cDNAs were amplified by PCR, turbo cloned (23) and sequenced. In order to amplify cDNA inserts from λ vectors, a single plaque was picked into 25 µl of distilled water, 1–5 µl were then added to a PCR and the cDNA insert amplified using vector-based primers. λgt10-specific primers, acaaatccacgtgttaatt and ggccctctttgggtatt (annealing at 68°C) and λgt11-specific primers gagagacatgccgtgaagtctgcc and gcacagcactcttgaagctagc (annealing at 56°C) were used to amplify inserts from the fetal brain and fetal heart cDNA libraries, respectively. Due to the probable introduction of sequence alterations during PCR, several subclones were sequenced. Alternatively, λ DNA was digested with EcoRI to release the cDNA insert which was then subcloned into EcoRI-digested pBluescript SK(–) (Stratagene).

DNA sequencing

In order to sequence the derived 11 PCR product, it was excised from a 0.8% low melting point agarose gel, dialysed in distilled water for 30 min, melted at 65°C and an appropriate quantity of melted gel slice added to the sequencing reaction.
Direct cosmid sequencing utilized 0.5–1 μg of cosmid DNA, with 60 ng of primer and ABI PRISM BigDye terminator cycle sequencing ready reaction kits (PE Applied Biosystems, Warrington, UK). Plasmid DNA sequencing reactions were performed using ABI PRISM dye terminator or dRhodamine terminator cycle sequencing ready reaction kits and 20 ng of primer. Products were separated on an ABI 373 or 377 DNA sequencer (PE Applied Biosystems), according to the manufacturer’s instructions. Resulting sequence was analysed using the GCG package of sequence analysis software (Wisconsin package version 9.1, Genetics Computer Group, Madison, WI). Sequence contigs were constructed using the Phred, Phrap and Consed software, version 6.0 (24–26).

BLASTN (27) searches using chromosome 11 breakpoint sequence identified only sequence matches to two bacterial artificial chromosome (BAC) end clones (accession nos AQ748746 and AQ105798). Searches using chromosome 1 breakpoint sequence identified sequence matches to a triplet repeat DISJ621 (accession nos G09671, G09453 and G07779), three BAC end clones (accession nos AQ112950, AQ784948 and B40542) as well as the mRNA and EST matches.

**Sequencing of cosmid ICRFc112I0142Q6**

This cosmid spans the chromosome 1 breakpoint (J.K. Millar, S. Christie, D. Lawson, D. Hsiao-Wei Loh, B. Arveiler and D.J. Porteus, manuscript in preparation). It was obtained by screening a chromosome 1 cosmid library provided by the Resource Centre/Primary Database (RZPD) of the German Human Genome Project at the Max-Planck-Institute for Molecular Genetics (http://www.rzpd.de). The probe used was the 2.7 kb derived 1 fragment. Cosmid DNA was treated with Plasmid-Safe ATP-dependent DNase (Epicentre Technologies, CAMBIO, Cambridge, UK) to remove contaminating *E.coli* DNA according to the manufacturer’s instructions. Cosmid DNA was partially digested with Sau3AI, and the resulting restriction fragments size-fractionated on a 0.8% agarose gel. Fragments of ~900 bp were excised from the gel and subcloned into pBluescript SK– (Stratagene). Subclones were picked randomly and sequenced using vector-based primers flanking the cloning site (caggaaacagctatgac and gtaaaacgacggccagt). Contig overlaps were established by designing primers from the ends of the contigs and subsequent direct cosmid sequencing. This process was repeated to generate the sequence of ICRFc112I0142Q6 as two ordered contigs (GenBank accession no. AF222987).

**PCRs**

Unless otherwise stated, PCR was carried out using AmpliTaq DNA polymerase (Perkin Elmer Biosystems, Foster City, CA). Each 50 μl reaction contained 1 U of enzyme, 300 ng of each primer, 200 mM of each dNTP, 1.5 mM MgCl₂, 50 mM KCl and 10 mM Tris–HCl pH 8.3. All reactions utilized 35 cycles with a denaturation step of 30 s at 94°C, an annealing step of 1 min at a temperature appropriate for the primers used, and a synthesis step at 72°C, based on the assumption that 1 min is required to synthesize 1 kb of DNA.

The 1.4 kb derived 11 breakpoint PCR product was amplified (annealing at 60°C) using one primer specific for chromosome 11 (ggctggatatgcctgtgaactaat) and one primer specific for chromosome 1 (agaacagaggagagatggac). M13S9 genomic DNA was used as template.

**Analysis of DISC1 expression by RT–PCR**

PCR was performed using proximal primers ggaaggagagggagccagggga and gacgctggactggttaagcaact (152 bp product) with the Advantage- GC cDNA PCR kit (Clontech) and cycling conditions as recommended by the manufacturer. Distal primers ggaaggagtgacggttcattace and agactcctcagctgtggttgac designed from DISC1 were also utilized for RT–PCR (270 and 336 bp products, annealing at 64°C). Analysis of DISC2 expression by RT–PCR utilized proximal primers ccacagaaccttcctcagataa and atcagcagaaatgtcagcgcgtg (250 bp product, annealing at 65°C) and distal primers gagagcagcagactgctgag and getcagcataagagactgtgac (486 bp product, annealing at 68°C). In the latter three cases, an initial hot start step was carried out.

**Hybridization**

Standard procedures were used for Southern blotting and hybridization (22). Double-stranded probes were labelled with [α-³²P]dCTP by random priming using High Prime (Boehringer Mannheim, Lewes, UK) and purified using Amersham Pharmacia Biotech (Little Chalfont, UK) NICK columns. The oligonucleotide probe was labelled with [γ-³²P]dATP using T4 polynucleotide kinase and hybridized to filters at 58°C.

**Sequencing of DISC1**

A 694 nt probe containing the 189 nt exon of AB007926 (contained within the 7.3 kb chromosome 1 breakpoint fragment) was prepared by PCR using the chromosome 1 breakpoint fragment as template and primers ccctgtaggctaatgctg and gacaactgtaagcctcactg. This probe and subsequent rounds of cDNA library rescreening generated several overlapping clones from the fetal brain and fetal heart cDNA libraries spanning 6913 nt of the transcript. 5′ RACE products were obtained from DISC1 using the Advantage-GC cDNA PCR kit (Clontech) and cycling conditions as recommended by the manufacturer, with human 20–25 week fetal brain Marathon-Ready cDNA (Clontech) as template (gene-specific primer 1, gactgtagctcactgtgct; gene-specific primer 2, gcagctgtgagctgtaac). Genomic sequence (2.4 kb) encompassing the 5′ end of DISC1 was obtained as follows. An oligonucleotide (ggaaggagagggagccagggga) designed from the 5′ end of DISC1 was hybridized to a genomic Psrl fragment of ~2.4 kb. Psrl fragments of this size were isolated from a PAC, 135-G6, containing the 5′ end of DISC1, (J.K. Millar, S. Christie, D. Lawson, D. Hsiao-Wei Loh, B. Arveiler and D.J. Porteus, manuscript in preparation), subcloned into pBluescript SK– (Stratagene) using standard methods (22) and sequenced (GenBank accession no. AF222982).

**Identification of DISC2 transcribed sequence**

Probes corresponding to ESTs AA249072 and W04811 were used to screen fetal heart and fetal brain cDNA libraries. A 557 nt probe for EST AA249072 was prepared using the chromosome 1 breakpoint fragment as template and primers gtgtcataatcagtagatggc and ctcctcgaacagggagtgtcc.
Northern blot analysis
A probe corresponding to nucleotides 958–1983 of DISC1 was obtained by excising one of the DISC1 fetal brain library cDNA clones using EcoRI. The cDNA clone corresponding to EST W04811 was used to identify DISC2 transcripts. Northern blots were obtained from Clontech.

ACKNOWLEDGEMENTS
We thank Kathy Evans for useful discussion and critical reading of this manuscript. This work was supported by Organon NV and the UK Medical Research Council.

REFERENCES

Human 3′ UTR length analysis
3′ UTRs were determined for the KIAA subset of mRNA sequences, based on coding sequence annotations. 3′ ends of the determined UTRs were checked for the presence of a consensus poly(A) signal (AATAAA) within 400 nt of the presumed end. UTRs failing to meet this criterion were excluded from further analysis. Sequence coordinates for full-length 3′ UTRs were determined and poly(A) signals detected using in-house software.


