Human Molecular Genetics, 2008, Vol. 17, No. 15 **2293–2309**
doi:10.1093/hmg/ddn130
Advance Access published on April 18, 2008

Variants in the estrogen receptor alpha gene and its mRNA contribute to risk for schizophrenia

Cynthia Shannon Weickert[1,4,†*], Ana L. Miranda-Angulo[1,†], Jenny Wong[4], William R. Perlman[1], Sarah E. Ward[1], Vakkalanka Radhakrishna[3], Richard E. Straub[3], Daniel R. Weinberger[3] and Joel E. Kleinman[2]

[1]MiNDS Unit, Section on Neuropathology, [2]Section on Neuropathology, CBDB and [3]Clinical Brain Disorders Branch, GCAP, NIMH, NIH, Bethesda, MD 20892, USA and [4]Schizophrenia Research Laboratory, Schizophrenia Research Institute, POWMRI, UNSW, Sydney, NSW 2031, Australia

Received January 3, 2008; Revised March 12, 2008; Accepted April 15, 2008

Estrogen modifies human emotion and cognition and impacts symptoms of schizophrenia. We hypothesized that the variation in the estrogen receptor alpha (ESR1) gene and cortical ESR1 mRNA is associated with schizophrenia. In a small case–control genetic association analysis of postmortem brain tissue, genotype CC (rs2234693) and haplotypes containing the C allele of a single-nucleotide polymorphism (SNP) in intron1 (PvuII) were more frequent in African American schizophrenics ($P = 0.01–0.001$). In a follow-up family-based association analysis, we found overtransmission of PvuII allele C and a PvuII C-containing haplotype ($P = 0.01–0.03$) to African American and Caucasian patients with schizophrenia. Schizophrenics with the 'at risk' PvuII genotype had lower ESR1 mRNA levels in the frontal cortex. Eighteen ESR1 splice variants and decreased frequencies of the wild-type ESR1 mRNA were detected in schizophrenia. In one patient, a unique ESR1 transcript with a genomic insert encoding a premature stop codon and a truncated ESR1 protein lacking most of the estrogen binding domain was the only transcript detected. Using a luciferase assay, we found that mRNA encoding a truncated ESR1 significantly attenuates gene expression at estrogen-response elements demonstrating a dominant negative function. An intron 6 SNP [rs2273207(G)] was associated with an ESR1 splice variant missing exon seven. The T allele of another intron 6 SNP was part of a 3′ haplotype less common in schizophrenia [rs2273206(T), rs2273207(G), rs2228480(G)]. Thus, the variation in the ESR1 gene is associated with schizophrenia and the mechanism of this association may involve alternative gene regulation and transcript processing.

INTRODUCTION

Schizophrenia, like other psychiatric disorders, has a strong genetic component. One approach to delineating 'causative' genes involves analyzing candidates with neurobiological plausibility in relationship to known characteristics of the disease, termed 'functional candidate association studies' (1). Many convergent lines of inquiry, from clinical observations, to measures of brain physiology, and molecular pathology suggest that estrogen signaling may be altered in the brains of patients with mental illness. First, the onset of disorders such as schizophrenia frequently occurs during or just after puberty (2,3) suggesting that sex steroid-triggered maturational changes in the brain may unmask vulnerability (4,5). Additionally, gender differences in symptoms (6,7), age of onset and course of disease over time have been well documented (8,9). Changes in schizophrenic symptoms are common during

*To whom correspondence should be addressed at: Maquarie Group Foundation Chair of Schizophrenia Research, UNSW Department of Psychiatry, Prince of Wales Medical Research Institute, Corner of Barker and Easy Street, Randwick, Sydney NSW 2031, Australia. Tel: +61 293991117; Fax: +61 293991005; Email: cyndi@powmri.edu.au or c.shannonweickert@unsw.edu.au
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

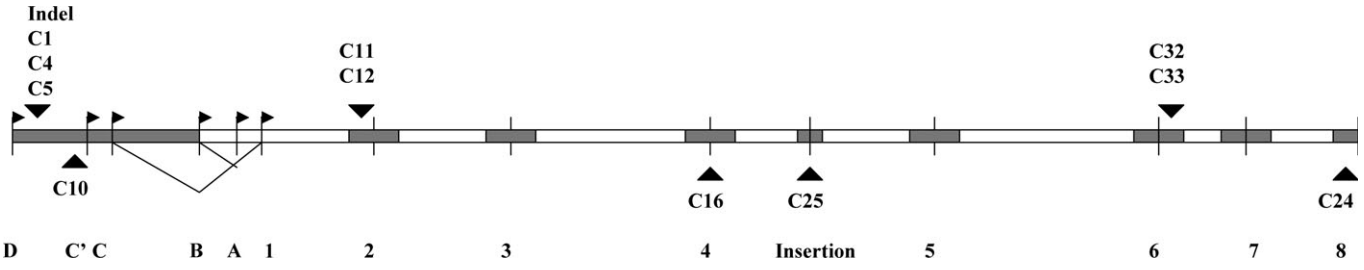

Figure 1. A schematic figure shows the general location of ESR1 common SNPs (arrows) used in association analyses. Gray horizontal bars indicate the regions of the ESR1 gene that were re-sequenced. The eight coding exons and the insertion are indicated by thin black vertical lines. ESR1 introns are shown as open horizontal bars (note thy can be grey or white). Numbers at the bottom correspond to exons and letters represent the different alternative start sites of transcription (indicated on gene with flags). Lines connecting exons C and B to exon 1 represent some of the alternative splicing. Other alternative 5′-UTR exons upstream of exon D are not shown.

pregnancy, at different phases of the menstrual cycle and during other periods of fluctuating hormonal levels (9–18).

Estrogen regulates emotional responses (19) and frontal cortical activity during cognitive task performance in humans (20), both of which are altered in major mental illness. While the transcriptional effects of estrogen are known to be mainly mediated by two gene products—estrogen receptor alpha (ESR1) (21) and estrogen receptor beta (ESR2) (22)—we have focused our initial efforts on ESR1 as we have previously found reductions in ESR1 mRNA in cortical and subcortical brain regions in patients suffering from major mental illness including a reduction in hippocampal ESR1 mRNA in patients with schizophrenia (23,24). Furthermore, telencephalic ESR1 mRNA levels significantly correlate with age of onset of schizophrenia and of major depression (23).

ESR1 is located on chromosome 6q25.1, spans over 300 kb, yields a prototypical mRNA of about 6.5 kb, contains eight coding exons and undergoes complex transcriptional regulation (25,26). Six alternative promoters regulate ESR1 gene expression in distinct organs, including brain, with promoters A, B and C utilized in human hypothalamus, amygdala, hippocampus and temporal cortex (27). Alternatively spliced ESR1 transcripts exist and are functionally important as ESR1 is a modular protein whereby removal of exons translates into stable proteins with differing ability to bind estrogen, translocate to the nucleus and interact with DNA (28). In frontal cortex of normal individuals, we have shown that alternative splicing of the ESR1 primary transcript can result in 12 distinct splice variants (29). The complement of ESR1 splice variants in frontal cortex differs between individuals, and we hypothesized that the variability in ESR1 mRNA isoforms may relate to psychiatric diagnosis and may be influenced by genetic variation in intronic regions proximal to splice donor and acceptor sites.

In this study, we hypothesized that the brain response to circulating estrogen may be altered in schizophrenia due to inheritance of particular forms of the estrogen receptor gene and/or due to the variation in estrogen receptor mRNA. We tested this possibility by determining if polymorphisms in the ESR1 gene are associated with the diagnosis of schizophrenia and whether differences in the level or type of cortical ESR1 transcripts are found in patients with schizophrenia compared to controls. We employed targeted re-sequencing of functional regions of the ESR1 gene in patients with schizophrenia to identify rare genetic alterations that may be specific to the disorder. We found associations between a functional SNP in intron one of ESR1 and the diagnosis of schizophrenia using our case–control brain cohort and confirmed this in a family-based association cohort. Additionally, we found that the 'at risk' SNP was associated with altered ESR1 mRNA levels in the frontal cortex of schizophrenics. We determined that as many as 20% of the patients with schizophrenia did not have detectable full length, fully functional, ESR1 mRNA in the frontal cortex and one patient produced a novel form of ESR1 mRNA that introduced a stop codon before the estrogen binding domain coding region. Furthermore, we found that the sequence variation in intron 6 predicts the presence of the Δ7 ESR1 splice variant in brain and is part of a 3-SNP haplotype associated with schizophrenia. Taken together, our results demonstrate that the ESR1 gene is a candidate to add to the growing list of potential 'susceptibility' genes for schizophrenia and that the mechanism of action includes changes at the mRNA level.

RESULTS

SNP identification by targeted re-sequencing

We successfully sequenced 6232 bp out of the 7787 bp targeted for re-sequencing in all individuals. The 1555 bp not re-sequenced corresponded to the region upstream of promoter A, the first exon and the 3′ flanking region of exon 1 (Fig. 1, Table 1) and were not able to be sequenced because of high GC content. Within the sequenced regions, we identified 44 SNPs. Eleven had very low frequency ($<1\%$) and showed questionable sequence quality, and thus are not reported in Table 1. Many SNPs, some of which were novel, with high sequence quality but with frequencies $\sim 1\%$ were found in this sample and were more common in African American patients (Table 1). Among the novel gene polymorphisms, we found an insertion/deletion polymorphism (indel) [-3815_-3816delAA] located at 554 bp upstream from the alternative transcription start site of exon C large that was subsequently submitted to NCBI by another group. The indel was followed by a G to A polymorphism (C1) which corresponds to rs6903180 in dbSNP. We found that the presence of the allele with the insert coincided in all cases with the presence of a G at the C1 SNP locus. Using direct re-sequencing from our family-based sample, we determined that the ESR1 promoter Indel was not transmitted in a

Table 1. Nomenclature for the SNPs used is defined in the far left column. The base pair substitution, the gene region, the rs number and the amino acid changes are given. The locations of the SNP defined in terms of the location on the chromosome, the intermarker distance and distance from the first ATG according to NM_000125 are computed. The minor allele frequency (MAF) in both racial groups is given for each SNP. The last four columns indicate the diagnosis of the individuals where rare SNPs (<0.06% frequency) were identified

										African american (n=64), Caucasian (n=92)[f]			
SNP name[a]	SNP[b]	Gene region	rs number	aa change[c]	SNP location UCSC freeze March 2006 (dbSNP build 126)[d]	Intermarker distance	Distance from ATG	African american MAF[e]	Caucasian MAF	Normal (n=18, 20)	Schizophrenic (n=45,43),	Bipolar (n=1,14)	Depresive (n=0,15)
INDEL	-3815_-3816delAA plus G/A		rs35903830		152166924	0	-3816	0.40	0.43				
C1		Region between exon D-C large	rs6903180/rs34535804		152166926	2	-3814	0.05	0.02				
C3	A/C		rs34392831		152167008	82	-3732	0.00	0.01		0,1		
C4	C/T		rs488133		152167137	129	-3603	0.08	0.31				
C5	T/C		rs9478245		152167280	143	-3460	0.15	0.04				0,1
		TSS			**152167440**		-3300						
C37	A/G		Novel		152167492	212	-3248	0.00	0.01		0,1		
C38	C/A		Novel		152167623	131	-3117	0.01	0.00		1,0		
C6	G/A		Novel		152167689	66	-3051	0.00	0.01		0,1		
C7	C/T	5' UTR exon C large	rs11963577		152167750	61	-2990	0.04	0.01				
C8	G/C		Novel		152168087	337	-2653	0.02	0.01		3,0	0,1	
C40	G/T		Novel		152168304	217	-2436	0.06	0.00				
C41	C/T		Novel		152168389	85	-2351	0.00	0.01		0,1		
C10	T/G		rs2071454		152168517	128	-2223	0.38	0.15				
					152168529		-2211						
		TSS			**152168530**		-2210						
None		5' UTR exon C Short											
					152168648		-2092						
C29	C/A	Region between B-C	rs9340770		152169795	1278	-945	0.01	0.13				
C47	G/A		Novel		152170060	265	-680	0.01	0.00	1,0			
		TSS			**152170186**		-554						
None		5' UTR exon B											
					152170337		-403						
Not sequenced		Region between A-B			152170740								
		TSS			**152170507**		-233						
Not sequenced		5' UTR exon A											
		Common splice site			**152170670**		-70						
		ATG			**152170741**		0						
Not sequenced		Exon 1											
					152171191		451						
C11-PvuII	T/C		rs2234693		152205028	34968	34288	0.45	0.49				
C12-XbaI	A/G	Intron 1	rs9340799		152205074	46	34334	0.25	0.34				
C14	G/T		rs33979243		152205225	151	34485	0.02	0.01	1,0	2,1		
					152205424		34684						
		Exon 2											
					152205614		34874						
C15	C/A	Intron 2	Novel		152243467	38242	72727	0.00	0.01			0,1	
					152243482		72742						
C30	C/T	Exon 3	rs4986934	Arg243Arg	152243568	101	72828	0.00	0.01				
					152243598		72858						
C31	G/A	Intron 3	rs9340846		152243684	116	72944	0.03	0.01				
					152307000		136260						
C16	C/G	Exon4	rs1801132	Pro325Pro	152307215	63531	136475	0.07	0.18				
					152307335		136595						
C27	C/T	Intron 4	rs35630684		152307351	136	136611	0.02	0.00		2,0		
None													
C25	G/A	Insertion	rs932477		152346288	38937	175548	0.11	0.11				
C36	T/C	Intron 4	Novel		152346461	173	175721	0.02	0.00		2,0		
C34	A/G		rs9340973		152374389	27928	203649	0.02	0.00		2,0		
					152374483		203743						
C35	C/T	Exon 5	Novel	Pro375Ser*	152374518	129	203778	0.00	0.01		0,1		
					152374621		203881						
C18	T/C	Intron 5	rs9340974		152374677	159	203937	0.03	0.00				
					152423818		253078						
None		Exon 6			152170740								
					152423951		253211						
C32	G/T		rs2273206		152424004	49327	253264	0.33	0.07				
C33	A/G	Intron 6	rs2273207		152424018	14	253278	0.06	0.07				
C21	G/A		Novel		152457205	33187	286465	0.01	0.00	1,0			
					152457212		286472						
None		Exon 7			152170740								
					152457395		286655						
None		Intron 7			152170740								
					152461559		290819						
C22	A/G		rs9341068		152461679	4474	290939	0.02	0.00		2,0		
C23	G/A	Exon 8	rs9341069		152461719	40	290979	0.00	0.01	1,0			
C24	G/A		rs2228480	Ser594Gly	152461788	69	291048	0.13	0.17				
					152461793		291053						

[a] Black boxes indicate SNP used in the association analysis

[b] Common allele listed first. *C1 Location assuming insertion allele at Indel locus

[c] Amino acid change based on transcript NM_000125. * Predicted amino acid change.

[d] UCSC=University of California Santa Cruz genome browser. Bold numbers indicate start and end of exons.

[e] MAF: minor allele frequency calculated from samples used for targeted resequencing, SNPs with frequency equal to zero are included because they had higher frequency in at least one the populations

[f] Diagnosis and race of individuals having novel rare SNPs (MAF=0.01-0.02)

Mendelian fashion in 9 out of the 122 families evaluated (7.37%). The deletion allele appeared *de novo* in 33% of the non-Mendelian transmissions (3/9 transmissions), whereas the insertion allele appeared *de novo* in 66% of the non-Mendelian transmissions (6/9 transmissions). These *de novo* polymorphisms were found in both African Americans and Caucasians.

In African Americans, only PvuII (C11) in intron one and the promoter Indel were in moderate LD (r^2); whereas most of the selected markers were not in high LD (Table 2). In Caucasians, moderate-high LD was found between three markers in the 5′ end of the gene: the Indel (promoter), C4 (promoter) and PvuII (C11, intron 1). Also, in Caucasians, one SNP in intron 6 (3′ end) of the gene (C33) was in

Table 2. The linkage disequilibrium results generated from the African American (upper) and Caucasians (lower) case–control genotypes are plotted for the INDEL and the ten SNPs used in this study. Note that the correlation coefficient r^2 is plotted above and that D′ is plotted below the diagonal

African american n=132

	INDEL	C1	C4	C5	C10	C11-PvuII	C16	C25	C32	C33	C24
INDEL		0.081	0.154	0.038	0.04	0.305	0	0	0.007	0.004	0.014
C1	1		0.007	0.13	0.058	0.065	0.007	0.007	0.003	0.005	0.012
C4	1	1		0.007	0.057	0.062	0.044	0.013	0.051	0.009	0.102
C5	0.485	0.513	0.668		0.068	0.001	0.019	0.015	0.056	0.003	0.031
C10	0.354	0.683	1	0.495		0.002	0.037	0.023	0.007	0.001	0.038
C11-PvuII	0.619	1	0.833	0.079	0.059		0.005	0.007	0.002	0.013	0.015
C16	0.019	1	0.226	1	0.715	0.237		0	0.006	0.055	0.091
C25	0.019	0.128	1	0.133	0.341	0.246	0.094		0.01	0.009	0.022
C32	0.088	0.181	1	0.464	0.087	0.07	0.288	0.36		0.139	0.044
C33	0.275	1	1	0.456	0.063	0.413	0.27	0.12	1		0.001
C24	0.219	1	0.448	1	0.584	0.268	0.378	1	0.635	0.222	

Caucasian n=95

	INDEL	C1	C4	C5	C10	C11-PvuII	C16	C25	C32	C33	C24
INDEL		0.029	0.657	0.012	0.017	0.318	0.047	0.02	0	0	0.005
C1	1		0.005	0.18	0.033	0.019	0.001	0	0.003	0.003	0.001
C4	1	0.308		0.002	0.063	0.319	0.082	0.011	0.06	0.01	0.01
C5	0.571	0.476	0.269		0.16	0.005	0.008	0	0.018	0.004	0.001
C10	0.443	0.387	1	0.691		0.061	0.008	0	0.015	0.01	0.006
C11-PvuII	0.706	1	0.892	0.305	0.615		0	0.001	0.001	0	0.004
C16	0.37	0.124	0.357	0.225	0.467	0.044		0.233	0.107	0.23	0
C25	0.313	0.231	0.175	0.203	0.002	0.058	0.642		0.171	0.435	0.006
C32	0.044	1	1	0.226	0.122	0.09	0.444	0.414		0.533	0.003
C33	0.058	1	0.47	0.863	0.836	0.056	0.755	0.794	0.915		0.001
C24	0.144	0.47	0.16	0.359	0.09	0.146	0.098	0.475	0.365	0.03	

Correlation coefficent r^2above and D' below diagonal

moderate LD with a SNP in intron 4 (C25) and was in moderate LD with a nearby SNP (+14 bp) in intron 6 (Table 2).

Genetic association analysis

For the African Americans controls derived from the postmortem brain cohort, all genotypes were in Hardy–Weinberg (H–W) equilibrium with the exception of the Indel. In contrast, not only the Indel but also the C4 and C11 genotypes of African American schizophrenics showed significant deviation from H–W equilibrium in this cohort. Due to the absence of H–W in the Indel genotypes in both cases and controls and the absence of Mendelian transmission in a significant percentage of families, this polymorphism was excluded from the association analysis. We detected significant genotype frequency (but not allelic) difference between African American schizophrenics and normal controls at C1, C5 and C11 (PvuII) (Table 3). However, the comparisons reaching significance for both C1 and C5 included comparisons to the homozygotes for the rare allele which did not contain any normal controls. The genotypic association to the common PvuII SNP was more robust and showed both a significant increase in the percentage of normals who were heterozygotes and a significant increase in the percentage of schizophrenics who were homozygote for the rare allele (CC genotype) (Table 3). In addition, we found that several three SNP haplotypes in the 5′ end of ESR1, one of which included the PvuII risk SNP, were significantly associated with schizophrenia using a fixed three SNP sliding window analysis (Table 4). When using various numbers of SNPs to form haplotypes with GRASP (2, 3, 4 and 5), the three SNP haplotypes were found to be the most informative (data not shown). For the Caucasian patients and controls derived from the postmortem brain cohort, all SNPs were in H–W equilibrium, but no significant association was found.

Next, we confirmed that the variation in the ESR1 gene was associated with schizophrenia using an independent cohort, the African American nuclear families from the NIMHGI sample. The PvuII-C (C11) allele was significantly overtransmitted to patients with schizophrenia (Table 5, $T = 23$, $NT = 14$, $P = 0.03$ global $P = 0.03$). When Caucasians families were added to the analysis, the overtransmission of PvuII-C allele was still present ($T = 60$, $NT = 41$, $P = 0.01$, global $P = 0.013$). Also, using the NIMHGI sample, the pairwise transmission disequilibrium test (TDT) analysis in Caucasians did not show overtransmission of any specific SNP; however, we did find overtransmission of one 3-SNP haplotype (C5-T,C10-T,C11-C; $T = 24$, $NT = 14$, $\chi^2 = 4.35$, global $P = 0.04$).

Promoter analysis via bioinformatics

The informatics analysis revealed that the ‘at risk’ haplotype of the ESR1 promoter was predicted to alter transcription factor binding (Fig. 2, Table 6). Analysis of the upstream region of

Table 3. Allelic and genotype case–control analysis in African Americans is shown

SNP	Alleles	Number of cases (frequency)				Number of controls (frequency)				Allelic association		Genotype association[a]		
		11 (%)	12 (%)	22 (%)	*n*	11(%)	12 (%)	22 (%)	*n*	x^2	*P*	Genotype comparison	x^2	*P*
C1[b]	G/A	35 (0.59)	2 (0.05)	2 (0.05)	39	67 (0.91)	6 (0.08)	0 (0)	73	1.22	0.27	II, III, IV	3.75, 3.69, 3,81	**0.05, 0.05, 0.05**
C4	C/T	35 (0.88)	3 (0.08)	2 (0.05)	40	70 (0.85)	10 (0.12)	2 (0.02)	82	0	0.96	n.s	n.s	n.s
C5[b]	T/C	26 (0.65)	12 (0.30)	2 (0.05)	40	64 (0.76)	20 (0.24)	0 (0)	84	2.86	0.09	II, IV	4.67, 4.27	**0.03, 0.04**
C10	T/G	17 (0.38)	22 (0.49)	6 (0.13)	45	36 (0.42)	34 (0.40)	16 (0.19)	86	0.01	0.93	n.s	n.s	n.s
C11[b]	T/C	16 (0.39)	12 (0.29)	13 (0.32)	41	23 (0.26)	48 (0.55)	16 (0.18)	87	0	0.96	I, III	5.15, 5.97	**0.02, 0.02**
C16	C/G	38 (0.86)	6 (0.14)	0 (0)	44	68 (0.79)	15 (0.17)	3 (0.03)	86	1.82	0.18	n.s	n.s	n.s
C25	G/A	35 (0.78)	10 (0.22)	0 (0)	45	66 (0.77)	19 (0.22)	1 (0.01)	86	0.07	0.79	n.s	n.s	n.s
C32	G/T	16 (0.39)	18 (0.44)	7 (0.17)	41	34 (0.41)	39 (0.47)	10 (0.12)	83	0.29	0.59	n.s	n.s	n.s
C33	A/G	41 (0.91)	4 (0.09)	0 (0)	45	70 (0.81)	15 (0.17)	1 (0.01)	86	2.37	0.12	n.s	n.s	n.s
C24	G/A	34 (0.76)	11 (0.24)	0 (0)	45	59 (0.69)	23 (0.27)	3 (0.04)	85	1.06	0.3	n.s	n.s	n.s

The 10 individual SNPs entered into the analysis are listed in the first column and the alleles defined as (1/2) are shown. Both the allelic association results and genotype associations for the distributions are given. Significant genotype associations are bolded. Chi-square and *P* for the significant genotype comparison. n.s, chi-square and *P*-value of no significant comparisons are not shown.
[a]Genotype comparisons: I, 11 versus 12; II, 11 versus 22; III, 22 versus 12; IV, (11+12) versus 22.
[b]Significant deviation from Hardy–Weinberg only in cases.

Table 4. Haplotypes across the ESR1 gene were constructed from adjacent markers and are depicted vertically with 1 representing the common allele and 2 representing the rare allele

Marker	Haplotypes[a]								
C1	2								
C4	1	1							
C5	2	2	2	2					
C10		2	2	1	1				
C11			2	2	1	2			
C16					2	2	1		
C25						1	1	1	
C32							2	2	2
C33								1	2
C24									1
Cases	0.08	0.20	0.10	0.00	0.04	0.02	0.38	0.37	0.01
Controls	0.01	0.05	0.00	0.05	0.06	0.05	0.28	0.22	0.09
P-value	**0.003**	**0.003**	**0.001**	0.06	0.21	0.30	0.18	0.04	**0.01**
Global *P*	**0.03**	**0.001**	**0.02**	0.00	0.85	0.40	0.50	0.07	**0.03**

A three-SNP sliding window haplotype analysis in the African American case–control sample showed that haplotypes containing rare SNPs in the 5′ end of the ESR1 were overrepresented in patients. Haplotypes containing rare SNP variants in the 3′ end of the gene were overrepresented in controls. Statistically significant haplotypes are bolded.
[a]Haplotypes with the lowest *P*-value and with a frequency higher than 0.03 in at least one of the groups.

Table 5. The ETDT analysis in the NIMHGI sample is shown

NIMHGI ETDT[a] SNP name	Overtransmitted allele	T[b]	NT[c]	*P*
INDEL	AA	2 (1)	0 (0)	0.09
C1	G	16 (0.61)	10 (0.38)	0.23
C4	T	17 (0.53)	15 (0.47)	0.72
C5	C	7 (0.54)	6 (0.46)	0.78
C10	T/G	12 (0.5)	12 (0.5)	1.00
C11-PvuII	C	39 (0.66)	20 (0.34)	**0.01**
C16	C	16 (0.64)	9 (0.36)	0.16
C25	G	12 (0.57)	9 (0.43)	0.51
C32	G/T	14 (0.5)	14 (0.5)	1.00
C33	A	16 (0.64)	9 (0.36)	0.16
C24	A	18 (0.56)	14 (0.44)	0.48

The individual ESR1 SNPs tested for transmission are shown in the first column and the overtransmitted allele is shown in the second. When the alleles were equally transmitted both alleles are shown. Statistically significant over transmission of PvuII C allele is bolded.
[a]ETDT only consider transmissions from informative parents.
[b]Number of transmitted alleles (frequency).
[c]Number of non-transmitted alleles (frequency).

ESR1 showed that the minor alleles of C1, C5 and C10 as well as the common allele of C4 introduced or disrupted the transcription factor binding site (TFBS) for myelin zinc finger transcription factor (Myt1), sex-determining *r*egion on the Y chromosome, *S*ry-related HMG b*ox*-5 (SOX5) and *n*erve growth *f*actor-induced protein *C* (NGF1C), respectively. In addition, the PvuII, C11 allele C was predicted to remove the TFBSs for activating enhancer binding protein 4 (AP-4) in the sense strand and the zinc finger protein ZNF238 (RP58) in the antisense strand. Among the SNPs analyzed for transcription factor binding, C10 and C11 were located in a highly conserved region across species based on the conservation track of UCSC genome browser (freeze May 2006). Also, these two SNPs are located within predicted promoters based on the El Dorado tool of Genomatix. These results suggest that C1, C4, C5, C10 and C11 might be functional SNPs influencing ESR1 expression regulation. However, these predictions need to be confirmed by transcription factor binding experiments as well as by determining the correlation between SNP genotype and ESR1 mRNA levels in patients.

Relationship of ESR1 genotype and diagnosis to ESR1 mRNA levels

First, we determined that RNA quality, as determined by Bioaglent RNA integrity number (RIN) value, did not differ

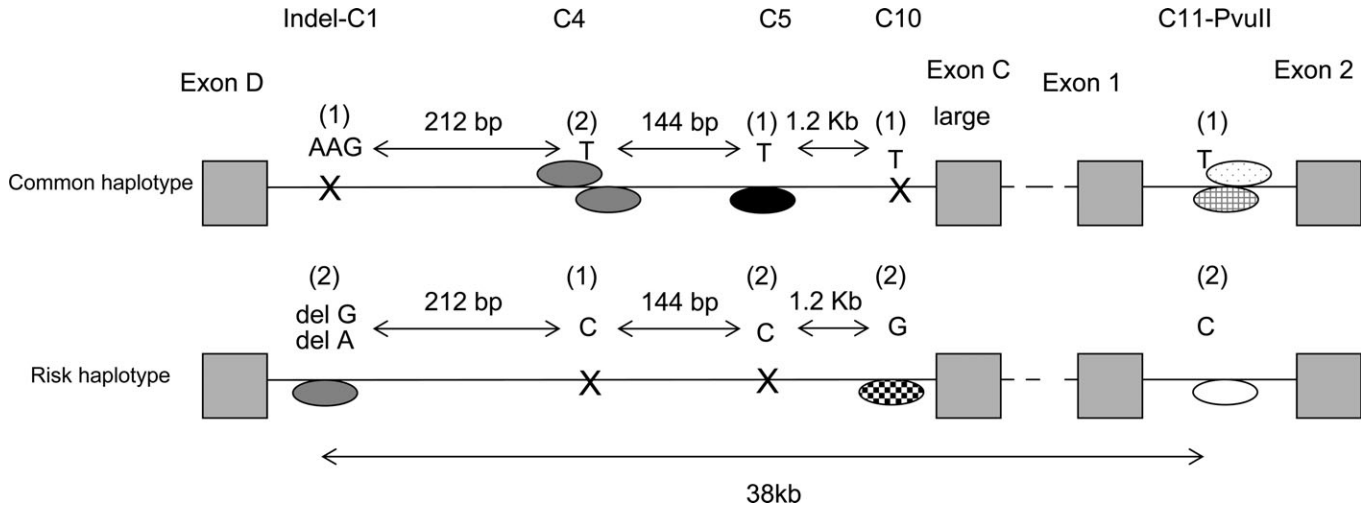

Figure 2. Schematic representations of the variation in transcription factor binding sites (TFBS) for the two promoter haplotypes of ESR1 are diagrammed. The top represents the common haplotype for C4, C5, C10 and C11 (1, 2, 1, 1, 1) that is not associated with the risk for schizophrenia. The bottom schematic shows the haplotype associated with risk (2, 1, 2, 2, 2). Ovals represent transcription factor binding to either the sense (upper) or antisense (lower) ESR1 DNA strand. Note the loss of a predicted biding of a transcription factor sites at C4, C5 and C11 and a gain of transcription factor binding site at C10 and Indel-C1 in the risk haplotype. Black oval, SOX5; grey ovals, MYT1; checkered oval, NGF1C; dotted oval, AP4; hatched oval, RP580 or ZNF238; white oval, myb.

Table 6. Transcription factor binding site analysis of identified ESR1 promoter SNPs

SNP-allele[a]	Family/matrix[b]	TFBS strand	Core similarity	Matrix optimized threshold	Matrix similarity	Core sequence[c]
indelC1-delC	V$MYT1/MYT1.01	−	0.75	0.75	0.76	aa***aAACTtcact c***
indelC1-delT	V$MYT1/MYT1.01	−	0.75	0.75	0.86	aa***aAACTtcact t***
indelC1-AAG	None	NA	NA	NA	NA	**NA**
C4-T	V$MYT1/MYT1.01	+	0.75	0.75	0.78	ag***aCAG T taacct***
C4-A	V$MYT1/MYT1.01	−	0.75	0.75	0.78	ag***aAGGTta a ctg***
C4-C	None	NA	NA	NA	NA	NA
C5-A	V$SORY/SOX5.01	−	1	0.87	0.99	gtc***a a CAAT*** ggcc
C5-C	None	NA	NA	NA	NA	NA
C10-T	None	NA	NA	NA	NA	NA
C10-C	V$EGRF/NGFIC.01	−	1	0.8	0.80	gaga***G C GTgtg t g*** g
C11-C	V$VMYB.05	–	1	0.9	0.95	taaAA**CGgctg**gg
C11-A	V$RP58/RP580.1	−	0.76	0.84	0.85	aa***aaC A GCtg*** gga
C11-T	V$AP4R/AP4.02	+	1	0.92	0.939	tgtcc***cAGC T g t***t ttat

The polymorphic change is given in the first column, the transcription factor identity is given in the second column. If no transcription factor was identified to bind to the interrogated sequence, none was entered in the second column and not available (NA) was entered into the subsequent columns. The third column shows the transcription factor binding site (TFBS) strand orientation where + equals the sense strand and where − equals the anti-sense strand. In the last column, the core binding site for the transcription factor within the 30 bp interrogated is bolded and italicized. The position of the SNP is underlined.

[a]SNP allele included in the input sequence for MatInspector analysis.

[b]None means that the SNP allele evaluated did not modify a TFBS, then the other columns information was not generated (NA).

[c]Core sequence in bold italic. SNP allele underlined.

significantly in patients with schizophrenia (RIN = 6.49) compared to controls (RIN = 6.16), ($t = 0.76$, df = 65, $P = 0.45$). Since DNA sequence variation in the upstream region of ESR1 may impact transcriptional rates of the ESR1 gene, we tested if polymorphisms within the putative promoter, intron 1 and 5′-UTR of ESR1 were associated with ESR1 mRNA levels in the DLPFC of patients with schizophrenia and/or controls. The PvuII SNP (C11), located in intron 1, showed no main effect of diagnosis ($F = 1.09$, $P = 0.301$, df = 1, 50) or genotype ($F = 0.774$, $P = 0.467$, df = 2, 50), but did show a significant interaction effect ($F = 4.54$, $P = 0.015$, df = 2, 50). A *post hoc* LSD test showed that schizophrenic patients with the C/C genotype (homozygous for 'risk allele') have significantly lower ESR1 mRNA levels than the patients who are homozygous for T/T (Fig. 3, $P = 0.025$). None of the other 5′ SNPs tested (Indel, C1, C4, C5 or C10) showed a significant main effect of genotype (all $F \leq 1.93$, $P \geq 0.16$) or an interaction effect (all $F \leq 2.21$, $P \geq 0.14$) on ESR1 mRNA level. Our attempts to relate haplotypes to mRNA levels were not successful because dividing individuals into groups based on haplotypes resulted in small numbers of subjects in certain haplotype groupings.

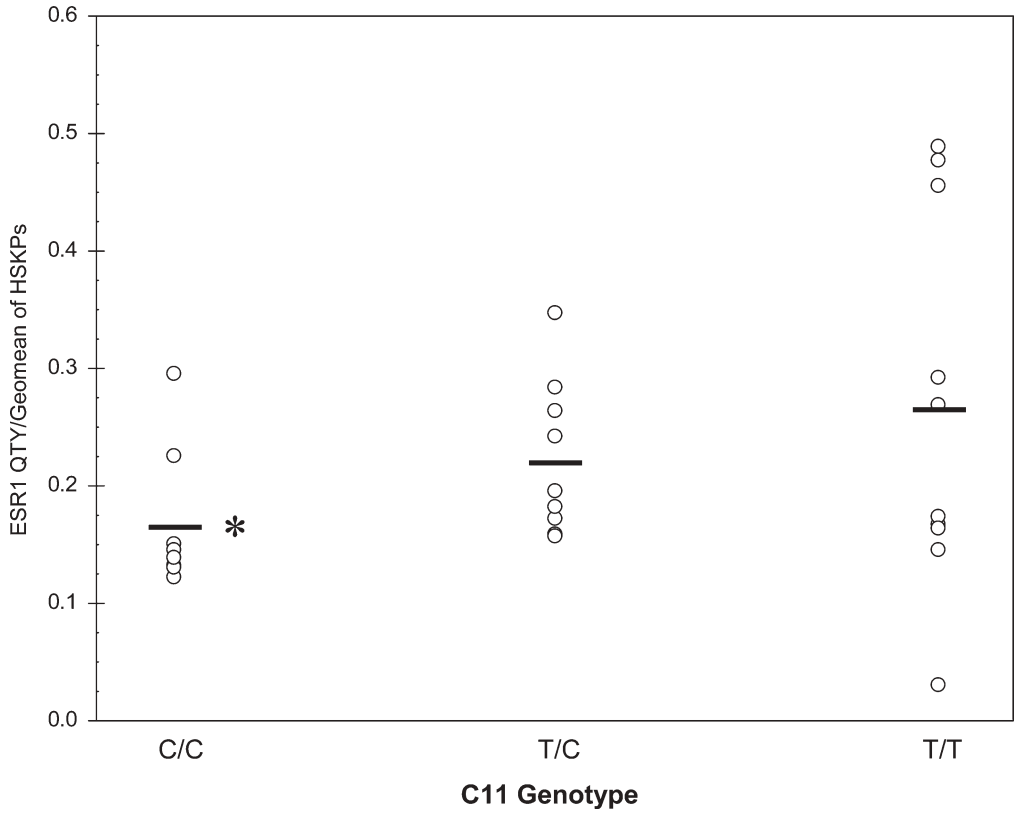

Figure 3. Scatter plot showing the average quantity of ESR1 mRNA levels normalized to the geometric mean of four housekeeper (HSKPs) mRNAs on the *Y*-axis as determined by real time qPCR. The values of ESR for patients with schizophrenia are plotted according to the PvuII (C11) genotype with C being the risk allele. Those patients homozygous for the risk allele (CC), have significantly lower expression compared to the other genotype groups (T/C and T/T). $^{*}P > 0.05$.

ESR1 splice variant detection

Following the first round of amplification of cDNA from the DLPFC, PCR products were often difficult to visualize on agarose gels (data not shown). After two rounds of PCR, a wild-type ESR1 mRNA of the expected size (1380 bp) (Fig. 4) was detected in 24 of 30 patients (80%) with schizophrenia, 13 of 15 patients with major depressive disorder (87%) and 12 of 15 patients with bipolar disorder (80%, Table 7). Other discrete PCR products were also detected in 47 of 60 patients (Fig. 4A, note multiple bands). Presumably, PCR products of smaller sizes represent exon-deleted ESR1 mRNAs; however, identification of which exon or exons were deleted based on the size of the product was not possible at this stage due to the similar sizes of some of the exons. In the third round of PCRs, where primers to specific splice junctions were used, we were able to detect a transcript population of 121 exon-deleted ESR1 mRNAs representing 18 distinct exon-deleted ESR1 mRNA transcripts from 60 human patient DLPFC samples using nested PCR (Fig. 5). We successfully amplified exon-deleted ESR1 mRNA variants using all our primer pairs on cDNA from the DLPFC of patients and their sequence was confirmed.

Detection of an ESR1 mRNA insertion variant in a schizophrenic patient

After two rounds of PCR amplification, one patient with schizophrenia with no detectable wild-type or exon-deleted ESR1 mRNAs was found to express an ESR1 mRNA larger than wild-type (1493 bp, Fig. 4A) with unpredicted sequence between exons 4 and 5. By using PCR primers targeted to exons 4 and 5, we detected 113 bp of sequence cleanly inserted between the last base of exon 4 and the first base of exon 5 (Fig. 4B). This sequence mapped to the ESR1 gene in the intron between exons 4 and 5 (Genbank accession no. AY425004.1 bases 177879–177991 in 6q25.1). The predicted protein coded by this variant mRNA would be 383 amino acids long with a stop codon at the end of exon 4 that would result in a truncated protein and absence of the estrogen binding domain.

Altered frequency of ESR1 transcripts in patients with major mental illness

Using the same methodology previously employed on the same brain region of 29 normal controls (29), we observed wild-type and 9 exon-deleted ESR1 mRNA variants in common with patients diagnosed with schizophrenia, bipolar disorder and major depressive disorder. Three ESR1 exon-deleted mRNA variants detected previously (Δ2,4,7, Δ6, and Δ7,2,3,4) were not observed in any of the 60 patients, although they had been found at only low frequency in normal controls (29). Six ESR1 exon-deleted mRNA variants not observed in normal controls (Δ3, Δ3,7, Δ5,2,3, Δ7,5, Δ7,2,3,4,5, and Δ7,2*3,4,5,6*) (with only partial deletions of exons 2 and 6)

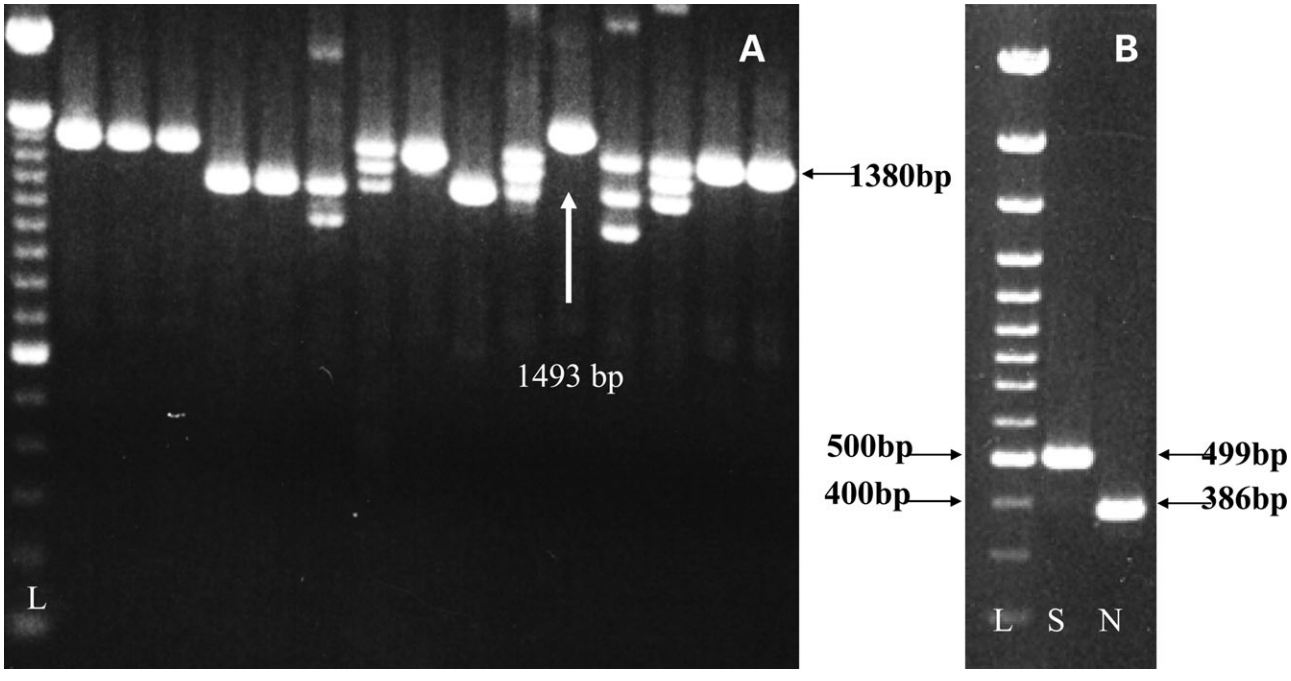

Figure 4. (**A**) The PCR products produced after nested PCR for ESR1 from cDNA derived from the DLPFC of patients with schizophrenia. Note that the samples are migrating on a slight downward diagonal. Lane 1 contains the 100 bp DNA ladder (L). The bright band in the first three sample lanes (lanes 2–4) and the last 2 lanes (15 and 16) corresponds to the expected 1380 bp amplicon (horizontal arrow). This ~1.4 band would be predicted if all 8 coding exons were present in the ESR1 transcript (wild-type). Many bands with smaller than the predicted sizes are found in the range of 1 to 1.3 kb and can occur in different individuals (for example, compare lanes 4 to lane 5) or within the same individual (lanes 7, 8, 11, 13 and 14). The exact identities of these bands could not be definitively identified at this stage (see Materials and Methods). Note that in lanes 5, 6, 7 and 10 no wild-type bands are detected. In the 12th lane, a larger than predicted amplicon was observed (vertical arrow, 1493 bp). In panel B, the 1493 bp ESR1 transcript was confirmed to be about 100 bp longer by a separate PCR directly targeting ESR1 exons 4 and 5 in normal individuals (N) and in the patient with schizophrenia with this genomic insertion (S).

were detected in patients with a mental illness, all at low frequency (Fig. 5, Table 7). We tested if the frequency of the two most common ESR1 transcripts (wild-type and Δ7 ESR1) differed in controls and patients with major mental illness, and we found that the frequency of wild-type ESR1 mRNA expression was lower in patients with schizophrenia (80%) when compared with normal controls (97%) ($\chi^2 = 3.86$, $df = 1$, $P < 0.05$) (Fig. 5, Table 7). We noted that there were also 20% of the patients with bipolar illness that lacked the wild-type ESR1 transcript; this did not, however, reach statistical significance likely due to the reduced power with the smaller number of subjects in the bipolar group. In the patients with major depression, we detected a higher frequency of the Δ7 ESR1 mRNA (80%) when compared with normal controls (48%) ($\chi^2 = 4.12$, $df = 1$, $P < 0.05$) (Table 7). Isoform variants were not found to vary significantly according to race or gender (29).

Δ7 ESR1 attenuates the wild-type ESR1 response to estrogen

Considering that patients with schizophrenia were more likely to express ESR1 splice variants, we determined if they could have a functional impact on wild-type ESR1 signaling. Since the Δ7 ESR1 transcript was the most abundant ESR1 splice variant in human brain and was associated with mental illness, we tested whether expression of the Δ7 ESR1 transcript would impact ESR1 mediated gene transcription. Using an estrogen-response element (ERE)-driven luciferase reporter construct containing three tandem repeats of a palindromic ERE, we first determined the ability of the wild-type ESR1 and Δ7 ESR1 to activate gene expression at different concentrations of 17β-estradiol (Fig. 6A). We showed that overexpression of wild-type ESR1 potently (about 3.6 times control vector) stimulated ESR1 mediated transcription in a concentration-dependent fashion (Fig. 6A), while overexpression of Δ7 ESR1 failed to elicit a transcriptional response at any dose. Next, we determined if the Δ7 ESR1 variant elicited a dominant negative effect on the wild-type ESR1 transcriptional response in the presence and absence of exogenous estrogen. We found a very significant difference in transcriptional activity in our four treatment conditions ($F = 1136.9$, $df = 3,20$, $P < 0.001$) (Fig. 6B). Importantly, when wild-type ESR1 was co-expressed with the Δ7 ESR1 mutant, the increase in wild-type ESR1-mediated transcription in the presence of 17β-estradiol was attenuated by over 50% ($P < 0.0001$). Additionally, a more modest but also statistically significant repression was found without the addition of exogenous estrogen, perhaps due to local estrogen synthesis by CHO cells in culture.

ESR1 polymorphisms and ESR1 exon-deleted mRNA variant detection

Since intronic ESR1 DNA sequence variation may impact on splicing of ESR1 transcripts, we asked if three distinct SNPs, C25, C32 or C33, all in introns and preceding one of the two

Table 7. All of the distinct ESR1 transcripts detected in the human prefrontal cortex are listed in the far left column

mRNA variant	Normal $n = 29$	Schizophrenia $n = 30$	Depression $n = 15$	Bipolar $n = 15$
WT (sign 1)	96.6	80.0	86.7	80.0
D 2	27.6	10.0	46.7	26.7
D 2,5	17.2	6.7	26.7	20.0
D 2,7	24.1	13.3	33.3	20.0
D 2,4,7	6.9	0.0	0.0	0.0
D 2,5,7	10.3	3.3	6.7	0.0
D 3	0.0	3.3	6.7	6.7
D 3,7	0.0	6.7	13.3	13.3
D 4	20.7	16.7	6.7	6.7
D 4,7	17.2	10.0	13.3	13.3
D 5	20.7	26.7	46.7	46.7
D 5,2,3	0.0	0.0	6.7	0.0
D 6	3.4	0.0	0.0	0.0
D 7 (sign 2)	48.3	40.0	80.0	53.3
D 7,5	0.0	6.7	6.7	0.0
D 7,3[a],4,5[a]	6.9	3.3	0.0	0.0
D 7,2,3,4	3.4	0.0	0.0	0.0
D 7,2,3,4,5	0.0	0.0	6.7	0.0
D 7,2[a],3,4,5,6[a]	0.0	0.0	6.7	0.0
INSERTION i4	0.0	3.3	0.0	0.0

WT, wild-type and D, delta followed by the missing exon number(s) are shown. The percentage of individuals in each diagnostic group where a particular ESR1 transcript was detected is given in the table. Note that since the same number of patients with bipolar and depression was evaluated, these percentages can sometimes be identical. Significant differences in frequency between the diagnostic groups compared to the normal control group are indicated with (sign). Significant difference ($P < 0.05$) between schizophrenic and normal controls (sign 1). Significant difference ($P < 0.05$) between depressives and normal controls (sign 2).
[a]Only partial exon is present in transcript.

commonly spliced exons (5 and 7) were associated with detection of the Δ5 or Δ7 ESR1 splice variants. Indeed, individuals whose ESR1 transcript pool contained specific exon 7-deleted mRNAs were more likely to have the common G allele at the intron 6 (C32 SNP) when compared to a T allele ($\chi^2 = 4.48$, $df = 1$, $P = 0.034$). The two other SNPs tested did not associate with the presence of either Δ5 or Δ7 containing transcripts. The C32 SNP was part of an associated haplotype (protective) detected at the 3′ end of the gene (Table 4).

DISCUSSION

ESR1 gene is associated with schizophrenia

In this study, we detected association of the estrogen receptor alpha gene with schizophrenia in two samples. Genotype (CC) of ESR1 SNP in intron 1, the PvuII (C11), was significantly associated with the diagnosis of schizophrenia in our African American case–control sample of postmortem brain tissue. Furthermore, SNP haplotypes located at the 5′ end of the gene and extending from the promoter into intron 1 were associated with the diagnosis of schizophrenia in this sample. The 3′ region of the ESR1 gene also showed some association when ESR1 haplotypes were analyzed. In an independent group of patients, we found further support for genetic association between the variation in the ESR1 gene and schizophrenia. The C allele of C11 found to be over-represented in patients with schizophrenia in the brain-derived case–control sample, was preferentially overtransmitted to offspring with schizophrenia using TDT analysis in the NIMHGI sample. Furthermore, the CC genotype at PuvII (C11) predicted lower ESR1 mRNA levels in patient prefrontal cortex. We found that the PvuII (intron 1) SNP was in moderate LD with a common insertion–deletion polymorphism (Indel) that we identified in the 5′ domain of the ESR1 gene (this Indel was deposited to NCBI by a different group during the course of our study). Additionally, we identified other rare ESR1 gene variants especially in patients. We observed significant deviation from H–W equilibrium in the genotype distribution of Indel, C1, C4 and C11 (PvuII) in patients with schizophrenia again suggesting an association between these polymorphisms and schizophrenia (30). It is unlikely that the deviations from H–W are due to genotype errors because many of these SNPs were confirmed by direct sequencing and the assigned genotypes were each verified manually. Also, the H–W deviation was not due to an excess of heterozygous individuals which is characteristic of genotyping errors. Taken together, our results suggest that DNA variants in the upstream region of the ESR1 gene may impart risk to developing schizophrenia and that the mechanism of this risk may involve altered transcriptional control of cortical ESR1 mRNA.

Genomic and mRNA changes in ESR1 implicate the promoter

Our computational bioinformatics analysis revealed that the variation at the 5′ end of the gene predicted altered transcription factor binding creating or deleting putative binding sites for five distinct transcription factors. However, two of these transcripts, Myt1 and SOX5, are expressed at fairly low levels in the adult (31,32), but are more abundantly expressed in the developing brain and have been linked to oligodendrocyte (Myt1) and cortical neuron development (SOX5) (33,34). AP4 and ZNF238 (also known as RP58) are expressed in cortical neurons (35,36). The most abundant of these potential ESR1 regulators is ZNF238 which is robustly expressed in the human prefrontal cortex and decreases significantly during postnatal life (data not shown). The NGF1C transcription factor (also known as EGR4) is expressed at a low to moderate level in the adult cortex and has been shown to be genetically linked to schizophrenia (37) and up-regulated by the antipsychotic, aripiprazole (38).

Changes in transcription factor binding could lead to reduced levels of ESR1 mRNA similar to what we have previously found in the brains of patients suffering from major mental illness (23,24). However, in our previous studies, genotype was not taken into consideration in the analysis. Direct experimental evidence shows the 1.3 kb intronic region of the ESR1 promoter containing the same polymorphic site which we have associated with schizophrenia [PvuII (C11)], is functional and changes transcriptional rate (39). In our study, we report evidence that patients homozygous for the C11 risk SNP (CC) have significantly reduced ESR1 mRNA in human frontal cortex. Indeed, we found that the presence of a C in the canonical CAGCTG site (with C11 underlined) of the gene was predicted to lead to a loss of AP-4 (transcriptional

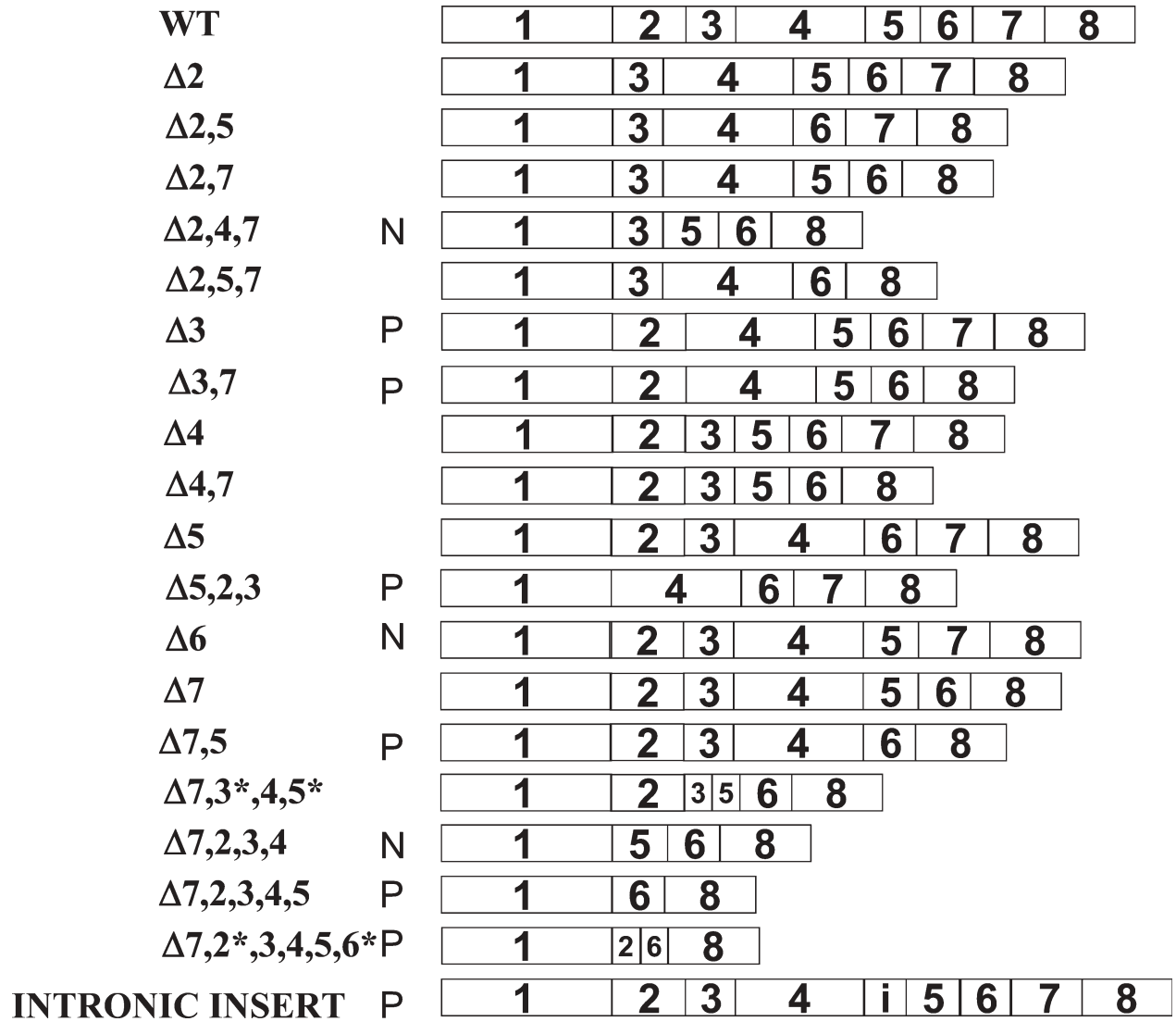

Figure 5. Schematic view of the ESR1 transcripts detected from cDNA from the DLPFC of adult humans. Nineteen different exon-deleted ESR1 mRNAs are depicted with the boxes representing the 8 known coding exons according to Genbank entry NM_000125. *the partial exon deletion, *N* the variant only detected in normal controls, *P* the variant only detected in psychiatric patients. WT = wild-type ESR1.

activator) binding on the sense strand and to loss of ZNF238 (transcriptional repressor) binding on the antisense strand, both of which would be expected to lower steady-state ESR1 mRNA (40,41). The other SNPs that are part of the 'at risk haplotype' containing C5 and C10 were also predicted to modify the binding of transcription factors SOX5 and NGFC1 to the upstream region of the alternatively spliced 5′ exon C.

ESR1 polymorphisms–disease comparison

Other studies suggest that the sequence variation in intron 1 of the human ESR1 may be associated with disease susceptibility as common SNPs in this genomic region, including PvuII, are linked to a variety of common human diseases, including breast cancer (42), osteoporosis (43,44), stroke (45), adiposity (46), atherosclerosis (47) and Alzheimer's disease (48). Furthermore, polymorphisms in intron 1 of ESR1 have been associated with risk of developing cognitive impairment in older women suggesting that these alleles may have functional consequences in the human brain (49). Despite widespread associations of ESR1 to many diseases including those impacting the brain, our study is the first to link common sequence variation in the estrogen receptor alpha gene to human brain ESR1 mRNA levels and to the risk of developing schizophrenia. However, it should be noted that there has been considerable evidence for linkage between loci on the long arm of chromosome 6, where ESR1 is located, and schizophrenia including a maximum multipoint LOD score of 7.7 on 6q25 in one of the largest pedigrees studied to date (50,51). Further support implicating the ESR1 locus as being involved in mental illness comes from one study that scanned about 5 kb of the ESR1 locus in 240 individuals with psychiatric disorders and identified four rare missense mutations in one bipolar, one alcoholic, one schizophrenic and one puerperal psychosis patient (52). In addition, the two common polymorphisms (Xba1 and PvuII) found in intron 1 of the

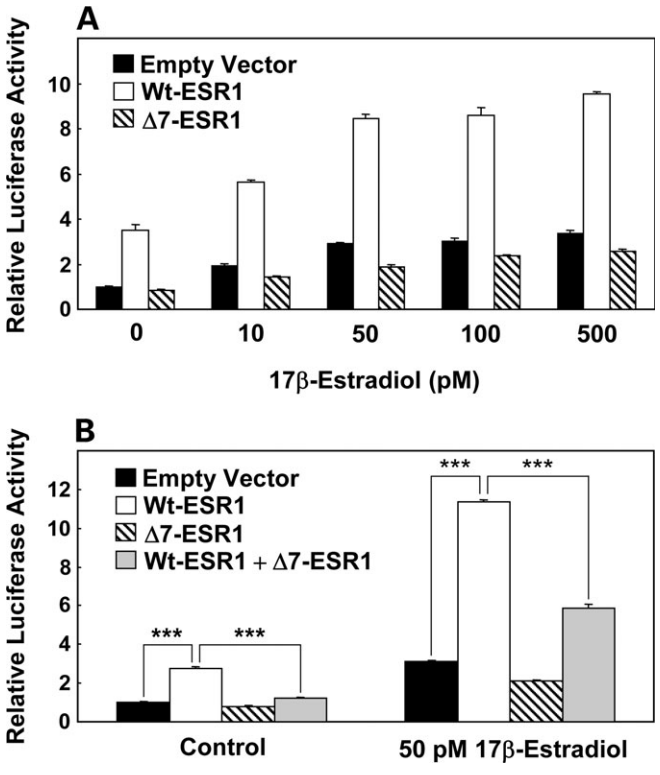

Figure 6. Cells were transiently transfected for 24 h with 3x ERE-luc and phRL-TK Renilla internal control plasmid together with either pDsRed-express-C1 (empty vector control), pDsRed-wt-ESR1 or pDsRed-Δ7-ESR1. Following transfection, cells were incubated for 24 h in the absence or presence of the indicated concentrations of 17β-estradiol (**A**) or 50 pM of 17β-estradiol (**B**). Values are mean +SEM averaged from $n = 6$ replicate cultures and presented as luciferase activity relative to the empty vector construct of the vehicle-treated control condition (***$P < 0.001$ by LSD or Tukeys).

ESR1 gene have been associated with the diagnosis of major depression in females, whereas no association with diagnosis of bipolar disorder or schizophrenia has been detected to date (53–56). The reasons for the discrepancy among our study and previous studies are not known but may relate to ethnic differences since our most positive evidence for association to ESR1 was found in African American individuals.

ESR1 splice variants are common in schizophrenia and in depressive patients

We found that patients with schizophrenia were significantly less likely to express the fully functional wild-type ESR1 transcript when compared with normal controls, with up to one-fifth of the patients with schizophrenia lacking any detectable wild-type ESR1 mRNA in their frontal cortex. The decreased frequency of the wild-type ESR1 in patients with schizophrenia was detected by one round of nested PCR; however, the various ESR1 splice variants were identified by two rounds of nested PCR and varied considerably from patient to patient. Since more than one nested procedure has substantial risk to amplify very minor or even 'junk' isoforms, the biological significance of these rare transcripts is not known. However, we did show that the expression of the commonly occurring Δ7 ESR1 splice variant was unable to elicit a transcriptional response at EREs, in line with previous studies which also reported a similar attenuation in ESR1 activity via various splice variants (28,57). Thus, we speculate that patients with schizophrenia may have a blunted molecular response to estrogen which may render the frontal cortex of many patients with schizophrenia unable to respond normally to brain estrogen. This could be biologically significant because changes in estrogen levels cause molecular and spine density changes in the prefrontal cortex (58–61). Additionally, treatment with estrogen can be an effective adjunctive therapy in some women with schizophrenia (62,63). One patient with schizophrenia synthesized only 'mutant' ESR1 transcript which contained an insertion of genomic DNA that leads to a truncated ESR1 protein missing the hormone binding domain rendering this individual's frontal cortex estrogen insensitive. Considering this, we would expect that some patients with schizophrenia may not be able to respond to 'estrogen therapy'.

Other patients, with major unipolar depression, were more likely to express the Δ7 ESR1 mRNA, compared to unaffected controls suggesting that unipolar depressed individuals may be more likely to express the dominant negative form of the ESR1 receptor which could alter the brain estrogen response. Indeed, we observed that the Δ7 ESR1 dampened the transcriptional activation mediated by wild-type ESR1 in response to estrogen in cultured cells. Additionally, circulating levels of estrogen are found to correlate with symptoms of major mental illness, especially with depression. Thus, alterations in sex hormone signaling in depressed patients could involve either altered circulating levels of sex steroids or altered endogenous brain response capabilities or both (29).

Rare genomic variations in ESR1 are common in patients with schizophrenia

After sequencing approximately 7 kb of genomic DNA in 162 individuals, we discovered novel SNPs in the ESR1 gene, most of which were rare and were typically detected in only one individual in our study. Patients with schizophrenia were more likely (13%) to have novel SNPs when compared with the non-schizophrenic individuals (6%). Interestingly, we identified six novel SNPs in the 5′-UTR of the transcript originating from exon C from six different patients with schizophrenia. The fact that we find many rare variants in ESR1 in schizophrenia is intriguing as there is interest in the possibility that multiple rare variants could underlie common diseases such as schizophrenia (50). Changes in the 5′-UTR of mRNAs can influence mRNA targeting and translation, change secondary mRNA structure and alter mRNA stability. For ESR1 mRNA, a recent study has found that transcripts with distinct 5′-UTRs have different translational efficiencies (64). Thus, it may be that in addition to (i) having altered frequency of genetic variation in a promoter of ESR1 that is link to reduced ESR1 mRNA levels, (ii) being less likely to splice together a fully functional wild-type ESR1 transcript, or (iii) synthesizing only a 'mutant' ESR1 transcript (our study and previous); patients with schizophrenia may be more likely to have rare genetic variation in the 5′-UTR ESR1 transcript. Each of these changes could independently or synergistically act to alter ESR1 protein production and ultimately, blunt estrogen function in brain. In this study, we have discovered several ways for ESR1 to differ in patients

Table 8. Demographics for all postmortem brains used in this study are given

Diagnosis	$n = 268$	Mean PMI	Mean age	Gender (n)	Mean pH	Race
Schizophrenia	95	36.88 ± 18.20	49.75 ± 16.95	M = 62, F = 33	6.36 ± 0.29	AA = 46, CA = 43, AS = 3, H = 3
Bipolar	15	32.53 ± 16.12	42.33 ± 11.7	M = 9, F = 6	6.18 ± 0.23	AA = 1, CA = 14
Depressive	15	27.46 ± 10.72	46.53 ± 9.31	M = 9, F = 6	5.76 ± 0.22	CA = 15
Control	143	29.14 ± 13.74	42.41 ± 15.19	M = 96, F = 47	6.48 ± 0.32	AA = 84, CA = 51, AS = 3, H = 5

The number of subjects in each group (n) and the average postmortem interval (PMI), age at death and tissue pH (+/− the standard error of the mean) are tabularized. M, males and F, females. AA, African Americans; C, Caucasians; AS, Asians; H, Hispanics. PMI, postmortem interval.

with schizophrenia compared to normals. We speculate that these genomic and transcriptomic differences in ESR1 could possibly lead to the same cellular fate, unresponsiveness to sex-steroid driven maturational brain changes during adolescence.

Summary

Since ESR1 is a complex genomic unit with transcription controlled by at least six promoters (65) and with at least 15 alternative splice variants of the coding domain, there is much opportunity for individual variation in how the ESR1 gene is expressed. Our data would suggest that the type and pattern of ESR1 transcripts may be altered in people suffering from schizophrenia. Our analysis suggests that the expression of ESR1 splice variants may be biased by genetic variation in the ESR1 gene. Furthermore, we present evidence that DNA variation in the 5′ region of the ESR1 gene may impact the susceptibility to developing schizophrenia and ESR1 mRNA levels in cortex. Taken together, our study suggests that ESR1 may be a 'candidate gene' for schizophrenia worthy of further investigation.

MATERIALS AND METHODS

In order to link ESR1 genomic changes with changes in ESR1 mRNA, the primary analysis of this paper was accomplished with postmortem brains of patients with schizophrenia compared to controls (Table 8, Supplementary Material, Cohort 1) and confirmatory genetic analysis was done on blood DNA from a sample of live patients and controls. Different subsets of individuals were used for the DNA re-sequencing (all cases Table 8) and case–control genetic association analysis from brain DNA (all Caucasians and African American normals and schizophrenics from Table 8), for the family-based genetic association from blood DNA (NIMHGI, Cohort 2), for the measurement of ESR1 mRNA via qPCR from brain RNA (Supplementary Material, Cohort 3), and for the ESR1 splice variant detection from brain RNA (Supplementary Material, Cohort 4). The decision of which samples to include in each of the various assays was based on availability of DNA, RNA and or cDNA at the time the experiment was initiated. We included patients with bipolar and depression as diagnostic contrasts to patients with schizophrenia for the ESR1 splice variant analysis based on a prior written agreement with the Stanley Foundation to include all of the diagnostic groups when using brains from their collection.

Postmortem diagnosis and exclusion criteria for samples from brain

Brains were collected in the Section on Neuropathology of the Clinical Brain Disorders Branch (CBDB) of the NIMH and the Stanley Foundation through the medical examiner offices following autopsy and through donations via funeral homes with consent from the next of kin. Psychiatric diagnosis was determined by telephone interviews with next-of-kin (including the Structured Clinical Interview for DSM-IV and the NIMH autopsy interview), and by review of all available clinical records and the Diagnostic Evaluation After Death (DEAD) (66,67). Cases that met DSM-IV criteria for schizophrenia, major depression and bipolar I disorder were used in our study. Disagreements between the two independent reviews were resolved by requesting an additional review by a third psychiatrist. Clinical information, neuroleptic dose calculations and brain tissue processing of the Stanley and NIMH cohorts have been previously described (68,69). Detailed inclusion and exclusion criteria for the schizophrenia cases used has been previously published (68,70).

Patient cohorts

Postmortem brains (from CBDB and Stanley Foundation). Two hundred and sixty-eight human brains of psychiatric patients and normal controls were obtained from the Clinical Brain Disorders Branch of the National Institute of Mental Health under an approved NIMH protocol (90M-0142) for human tissue and from the Stanley Medical Research Institute (Bethesda, MD, USA). This cohort included 143 non-psychiatric controls and 95 schizophrenics. Genomic DNA from 132 individuals (which included all but one patient with schizophrenia) from this cohort was sent to Polymorphic DNA Technologies, Inc. (Alameda, CA, USA) for targeted re-sequencing and SNP discovery (94 schizophrenics and 38 normals).

Clinical Sample from the National Institutes of Mental Health Genetic Initiative (NIMHGI). We used a subset of the NIMHGI samples consisting of 122 nuclear families and including one or two affected offspring (either schizophrenic and/or schizoaffective diagnosis) and at least one parent (51 African Americans families, 71 Caucasian families). This subset was constructed from the larger affected sib pair dataset to maximize the number of trios available for family-based association analyses. The diagnosis evaluation, collection and DNA isolation performed in the NIMHGI has been previously described (71).

Isolation of nucleic acid from brain

Genomic DNA was extracted from pulverized frozen brain (cerebellum or frontal pole) according to the protocol provided by the manufacturer of the extraction buffers (PureGene). Total RNA was extract from all cases from the dorsolateral prefrontal cortex (DLPFC) tissue dissected from the middle frontal gyrus. Frozen brain tissue was pulverized on dry ice and weighed while frozen prior to RNA extraction with the TRIZOL Reagent method (Life Technologies Inc., Grand Island, NY, USA) as previously described (72).

SNP detection by targeted re-sequencing and genotyping

We selected the following genomic regions to re-sequence in a search for novel DNA sequence variants related to transcriptional control or alternatively processed ESR1 transcripts: (i) 3.8 kb upstream of the ESR1 translation start site comprising the A, B and C promoters as defined by Kos *et al*. (25); (ii) the eight coding exons; (iii) a cassette exon in intron 4 (identified herein via RT–PCR); and (iv) 200 bp surrounding each exon of the ESR1 gene for a total of 7787 bp of brain DNA. Polymorphic DNA Technologies, Inc. (Alameda, CA, USA) performed the sequencing and SNP discovery by nucleotide comparison of single and bidirectional DNA sequences traces with the Paracel's *A*utomated *GEN*otyping *T*ool (AGENT) program. We verified the reported SNPs with Mutation surveyor program version 2.51 by visual inspection, sequence quality (>10) and mutation score (>20) which correspond to a confidence of 99% for correct call. To determine if a verified SNP was novel or known, the sequence around the SNP was submitted to the BLAT tool of University of California Santa Cruz (UCSC) genome browser on human assembly May 2006 and confirmed at dbSNP build 126. SNPs with a minor allele frequency (MAF) higher than 5% in at least one of the populations (Caucasians or African Americans) were selected to be genotyped by TaqMan® assays in 99 additional normal control brains (70% African Americans, 30% Caucasians) and in 122 NIMHGI nuclear families. Applied Biosystems' (ABI) TaqMan® primers and probes were constructed using either the ABI 'assays by design' or 'assays on demand' software and the assays were performed under standard conditions: 50°C for 2 min, 95°C for 10 min and then 40 cycles of 95°C for 30 s, 60°C for 1 min. The indel and C1 genotypes were obtained by re-sequencing the promoter region in these additional 99 controls (Polymorphic, Inc.).

Genotypes generated with TaqMan® assays were manually assigned using SDS 2.2.2 software with 10% sample redundancy to estimate genotyping accuracy (over 99%) and assays were repeated until over 98% of the genotypes were determined. Direct comparisons between genotypes obtained by direct sequencing and TaqMan® in a subset of 45 people, showed that 9/10 SNPs had over 98% consistency. However, due to high level of inconsistencies in genotypes of one SNP (rs9340799, XbaI), the data for this SNP were deemed unreliable and were dropped from all analyses.

Bioinformatics analysis–promoter analysis

A 30 nucleotide sequence surrounding each of the SNPs in the associated promoter haplotypes, C4,C5,C10 and the intron 1 SNP C11 were tested for introducing or disrupting a TFBS searching in the matrix family vertebrate library version 5.0 released on February 2005. We used the MatInspector tool of the Genomatix suite which uses position-specific weight matrices (PWMs) to calculate matrix similarity scores for the two versions of the sequence differing only in the SNP nucleotide. A good matrix similarity score is higher than 0.8 and a score of 1.0 is only reached when highest conserved bases of a matrix match exactly the target sequence. To reduce the number of false positives, the optimized matrix similarity score is also calculated. In a reliable prediction, the matrix similarity score is higher than the optimized matrix similarity score (73). The El Dorado tool of Genomatix was used to determine if any of the SNPs of interest were located within known or predicted promoters of the ESR1 gene.

qPCR from brain RNA

One hundred and forty-three brains of schizophrenic patients ($n = 53$) and normal controls ($n = 90$) were available for qPCR (subset of Table 8, see Supplementary Material information). Extensive clinical and molecular characterization of this cohort has been recently described (68). Some samples were excluded ($n = 21$) after a rigorous evaluation of RNA quality and factors influencing the expression of housekeeping genes (68). From this refined cohort, we chose a subset of 34 controls ($n = 17$ African American) and 34 schizophrenics ($n =$ 16 African American) matched for age, pH and PMI (see Supplementary Material tables), in order to examine any diagnostic difference in patients versus controls taking ESR1 genotype into consideration by two-way ANCOVA. We verified that the various genotype groupings formed did not significantly differ in subject age, brain pH, postmortem interval (PMI) or RNA integrating (RIN).

ESR1 mRNA expression levels were measured from RNA extracted from the dorsolateral prefrontal cortex. We synthesized first-strand cDNA using random primers and following the standard protocol outlined in the Invitrogen Superscript First-Strand Synthesis System (Invitrogen life technologies, Carlsbad, CA, USA). The qPCR assays were performed with TaqMan® technology on an ABI 7900 HT SDS machine (Applied Biosystems, Foster City, CA, USA). A probe with the sequence 5′-ATGATGAAAGGTGGGATACGAAAAG-3′ was designed to anneal to the junction of ESR1 exons 3 and 4 which was chosen to include the greatest number of exon deleted mRNA variants [Δ3 and Δ4 are among the least common deletions (29)]. Each qPCR reaction contained 3 μl cDNA template (corresponding to 90 ng cDNA), 5 μl Master Mix (Eurogentec, California), 0.5 μl 20X assay-on demand mix containing forward and reverse primers and FAM^{TM} dye-labeled TaqMan® MGB probe (Hs00174860_m1) and 1.5 μl H_2O. Both standards and samples were assayed in triplicate and gene expression levels were calculated as an average of the three. Samples were heated for 2 min at 50°C and 10 min at 95°C and then subjected to 40 cycles of denaturation at 95°C for 15 s and annealing and elongation at 60°C for 1 min. PCR data were acquired from the Sequence Detector Software (SDS version 2.2, Applied Biosystems) and quantified by the relative standard curve method using serial dilutions of pooled cDNA. SDS software plotted real-time fluorescence intensity

and selected the threshold within the exponential phase of the amplicon profiles. The software plotted a standard curve of the CT (cycle threshold) versus the RNA quantity. Triplicate measures were made and values with over a 3% standard deviation (CT) were omitted in calculating the averages per subject. ESR1 mRNA levels were normalized by dividing the ESR1 quantities by the geometric mean of the quantity of four housekeeper genes: SDHA (Hs01549169_m1), GUSB (Hs99999908_m1), PGDB (Hs00609297_m1) and cyclophilin A (Hs99999904_m1). No-template controls and no RT controls were included in the qPCR reaction plate and showed no detectable fluorescence.

ESR1 mRNA variant detection

Due to the multiple rounds of PCR, and the numerous gels and successive sequencing needed to determine the entire pool of ESR1 variants in each subject, we preformed a complete transcript ESR1 characterization on a restricted set of cases that was available when the experiments were initiated. Patients with bipolar disorder and depression were included ($n = 15$ and $n = 15$) in this assay. For diagnostic comparisons, patients with schizophrenia were matched to normal controls based on gender, race, brain pH and PMI. Demographic variables for the individuals used in the ESR1 variant analysis ($n = 89$) are described in supplementary information (Supplementary Material, Table S4). Three rounds (two nested) of PCR were used to amplify exon-deleted ESR1 mRNA variants from cDNA synthesized from frontal cortex as described above. Primers, cycle conditions, and expected fragment sizes for all three rounds of PCR have been previously described (29). In one schizophrenic case, after the first two rounds of PCR, a product 100 bp greater than expected size was amplified (see results, Fig. 5). Upon bidirectional sequencing, unpredicted sequence was found to exist between exons 4 and 5 (originating from intron 4, a cassette exon). To specifically examine this region of the mRNA, we conducted PCR on all cases using a forward primer designed against sequence in exon 4 (5′-acatgagagctgccaacctt-3′), and a reverse primer designed against sequence in exon 5 (5′-ggttcctgtccaagagcaag-3′). A 30 μl reaction mixture was assembled consisting of second round cDNA diluted 1:1000, dNTP mix (0.2 mM in final concentration), forward and reverse primers (0.2 μM each in final concentration), Platinum *Taq* PCR buffer, 1.5 mM $MgCl_2$ and 1.8 U of Platinum *Taq* DNA polymerase (Invitrogen life technologies, Carlsbad, CA, USA). The cycle conditions were 94°C for 30 s, 55°C for 30 s and 72°C for 1 min for 35 amplification cycles and the expected wild-type fragment size was 499 bp. PCR products were electrophoresed in 2% agarose gels, and selected products were excised, gel purified using the QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA), and bi-directionally sequenced (Lofstrand Labs Unlimited, Gaithersburg, MD, USA).

Generation of plasmid construct

pDsRed-wt-ESR1 and pDsRed-Δ7-ESR1 were constructed by PCR amplification of full-length human wild-type ESR1 or Δ7 ESR1 from commercially available MCF-7 cDNA (BioChain Institute, Inc. Hayward, CA, USA). Primer pairs used in the amplification contained both *Xho*I and *Bam*I restriction sites. Primer sequences were: Forward 5′-CTCGAGACCATGACCC TCCAC-3′ and Reverse 5′-GGATCCTCAGACCGTGGCA GGGA-3′. Each 25 μl reaction contained 0.5 μl cDNA template, PCR buffer, 0.2 mM dNTPs, 2 mM $MgCl_2$, and 0.02 U of KOD Hot Start DNA polymerase (Merck KGaA, Darmstadt, Germany). The cycling conditions were 95°C for 2 min, 40 cycles of 95°C for 40 s, 68°C for 1 min, 70°C for 2.5 min and 70°C for 3 min. PCR products were electrophoresed in 1% agarose gels and products corresponding to the approximate size of wild-type and Δ7 ESR1 were excised and gel-purified using the QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA). For the addition of 3′-A overhangs, a 25 μl reaction containing purified PCR product, PCR buffer, 0.2 mM dATP (Promega Corporation), 2 mM $MgCl_2$ and 0.5 U of Red Hot® DNA polymerase (ABgene, Epsom, Surrey, UK) were incubated at 72°C for 30 min. Reaction products were then purified and ligated into pcDNA3.1-V5-His using the TOPO TA cloning kit (Invitrogen Australia Pty Limited, Victoria, Australia) according to the manufacturer's instructions. Positive clones containing inserts for ESR1 wild-type and Δ7 were verified by sequencing. Both genes were then excised by restriction endonuclease digestion with *Xho*I and *Bam*I, gel purified and ligated into pDsRed-express-C1 (Clontech Laboratories, Inc. Mountain View, CA, USA) using T4 DNA ligase (Fermentas, Glen Burnie, MD, USA) according to the manufacturer's instructions.

Cell culture

CHOK-1 cells were grown at 37°C in a 5% CO_2 atmosphere and cultured in Dulbecco's modified Eagle's medium/Ham's F-12 medium (1:1 mixture) containing 10% (v/v) heat- inactivated fetal bovine serum supplemented with penicillin/streptomycin 100 U/100 μg/ml and glutamax (2 mM) (Invitrogen Australia Pty Limited, Victoria, Australia).

Luciferase reporter assay

On day 1, CHOK-1 cells were seeded into 48-well plates and grown to 50% confluence. On day 2, the growth media was removed and cells were washed twice in phosphate buffered saline. Cells were then refreshed with Dulbecco's modified Eagle's medium/Ham's F-12 medium (1:1 mixture) containing 10% (v/v) charcoal-stripped-heat-inactivated fetal bovine serum supplemented with penicillin/streptomycin 100 U/100 μg/ml and glutamax (2 mM) (Invitrogen Australia Pty Limited, Victoria, Australia). On day 3, CHOK-1 cells were transiently transfected for 24 h using Lipofectamine 2000 (Invitrogen Australia Pty Limited, Victoria, Australia) (0.5 μl/well) in charcoal-stripped media containing no antibiotics. Plasmids transfected included the 3x ERE-luc reporter plasmid (125 ng/well) (a kind gift from Professor Rakesh Kumar, PhD (M.D. Anderson, Houston, TX, USA), pDsRed-wt-ESR1 (12.5 ng/well) and/or pDsRed-Δ7-ESR1 (12.5 alone and 37.5 ng/well in combination with wt). The empty vector pDsRed-express-C1 was used to total the DNA content of the overexpression plasmids to 50 ng/well. The phRL-TK Renilla internal control plasmid (12.5 ng/well) was also co-transfected for normalization of transfection efficiency. On day 4, the transfection media was removed and

cells were incubated with the absence and presence of 50 pM of 17β-estradiol (Dr Ehrenstorfer GmbH, Augsburg, Germany) for 24 h. On day 5, cells were washed and resuspended in 50 μl of 1× passive lysis buffer (Promega Corporation). Luciferase assays were performed using the Dual Luciferase Assay Reporter System (Promega Corporation) according to the manufacturer's instructions in a FLUOstar OPTIMA plate reader (BMG Lab Tech, Offenburg, Germany). The results were normalized and expressed as the ratio of luciferase activity to the Renilla internal control, and normalized to the vehicle-treated control condition.

Statistical analysis

Case–control association analysis. Schizophrenics and controls from the postmortem brain collection were stratified by race to test for genetic association between ESR1 and schizophrenia within each racial group. Caucasians and African Americans of both genders were included in this case–control genetic analysis. Common SNPs ($N = 10$, MAF$>6\%$) of the two case–control samples corresponding to 88 schizophrenics (45 African Americans and 43 Caucasians) and 137 normal controls (87 African Americans and 50 Caucasians) were typed and used for pairwise and haplotype case–control association analysis. Pairwise case–control allele and genotype frequency comparison were performed with FINETTI version July 2002 (http://www.mds.qmw.ac.uk/statgen/dcurtis/software.html) and COCAPHASE of the UNPHASED program version 3.0.2 (74). Marker-to-marker LD for both Caucasians and African Americans was calculated using HAPLOVIEW version 3.32 (75).

Family-based association analysis. In order to test for association between ESR1 and schizophrenia in an independent clinical sample, we performed a TDT in DNA from the nuclear families available from the NIMHRGI. In these nuclear families, we genotyped the same high frequency SNPs which were genotyped by TaqMan and re-sequencing in the postmortem brain sample. Mendelization errors and improbable recombination were checked with MERLIN 0.10.2 and excluded from the analysis. We tested for allele and haplotype transmission disequilibrium using TDTPHASE of the UNPHASED program version 2.03. The allele and haplotype analyses were performed in each population (African Americans and Caucasians) separately and combined. In using TDTPHASE, we assumed no linkage and excluded missing data. We ran the pairwise analysis by including only certain transmissions. The uncertain option was used to increase the number of analyzed transmissions in the haplotype analysis.

Diagnosis and genotype analyses of prefrontal cortical ESR1 mRNA levels. ESR1 genotypic and diagnostic effects were analyzed in relation to ESR1 mRNA expression via two-way ANCOVAs. ESR1 mRNA levels decrease with age ($r = -0.413$, $P = 0.001$); therefore, age was used as a covariant in all statistical tests. PMI, pH, RIN and neuroleptic treatment were not correlated with ESR1 mRNA quantities. ESR1 mRNA levels were averaged from the triplicate measures and population outliers that were two standard deviations above or below the mean of the entire cohort were removed. ANCOVAs were run in two different ways with respect to the genotype variable. Genotype (11 versus 12 versus 22) was considered the independent factor and if there was a very low sample size ($n = <5$) in the homozygous rare allele group, the genotypes were combined into carriers of the rare allele (12 and 22) versus non-carriers of the rare allele (11) to increase the power. For all SNPs with a significant effect of diagnosis, genotype or an interaction effect, *post hoc* Fisher's LSD test was performed.

ESR1 mRNA variants frequency and SNP association analysis. Observed expression frequencies of the wild-type and common exon-deleted ESR1 mRNA variant were subjected to chi-square statistical analysis to determine if frequencies of ESR1 mRNA variant detection in patients diagnosed with major mental illness differed from the frequency of detection observed in normal controls (29). In addition, genotypes of SNPs with a MAF $>5\%$ located within 200 bp either in the intron preceding or following the ESR1 exons 5 (C25) and 7 (C32 and C33), were used to examine the potential relationship between the most common ESR1 mRNA deleted variants (Δ5 and Δ7) and the DNA variation present in putative splicing regions.

Luciferase reporter assay for ESR1 wild-type and Δ7. The effect of transfection of control, wild-type and wild-type +Δ7 ESR1 plasmids on ERE-driven luciferase reporter activity at two different dosages was assessed by two-way ANOVA (treatment condition and estrogen levels as IV) and subject to post-hoc Fisher's LSD analysis ($n = 6$ wells in each condition).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENT

Funding to pay the Open Access publication charges for this article was provided by, the Schizophrenia Research Institute, utilizing funding from NSW Health and the Macquarie Group Foundation.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the Intramural Program of the National Institute of Mental Health and by a National Alliance for Research on Schizophrenia and Depression Young Investigator Award to C.S.W. and by, the Schizophrenia Research Institute, utilizing funding from NSW Health and the Macquarie Group Foundation.

REFERENCES

1. Blair, I.P., Badenhop, R.F., Scimone, A., Moses, M.J., Kerr, N.P., Donald, J.A., Mitchell, P.B. and Schofield, P.R. (2005) Association analysis of transcripts from the bipolar susceptibility locus on chromosome 4q35, exclusion of a pathogenic role for eight positional candidate genes. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.*, **134**, 56–59.
2. Lieberman, J.A., Perkins, D., Belger, A., Chakos, M., Jarskog, F., Boteva, K. and Gilmore, J. (2001) The early stages of schizophrenia: speculations

on pathogenesis, pathophysiology, and therapeutic approaches. *Biol. Psychiatry*, **50**, 884–897.
3. Burke, K.C., Burke, J.D., Jr, Regier, D.A. and Rae, D.S. (1990) Age at onset of selected mental disorders in five community populations. *Arch. Gen. Psychiatry*, **47**, 511–518.
4. Styne, D.M. (1994) Physiology of puberty. *Horm. Res.*, **41** (Suppl. 2), 3–6.
5. Weinberger, D.R. (1987) Implications of normal brain development for the pathogenesis of schizophrenia. *Arch. Gen. Psychiatry*, **44**, 660–669.
6. Goldstein, J.M. and Link, B.G. (1988) Gender and the expression of schizophrenia. *J. Psychiatr. Res.*, **22**, 141–155.
7. Shtasel, D.L., Gur, R.E., Gallacher, F., Heimberg, C. and Gur, R.C. (1992) Gender differences in the clinical expression of schizophrenia. *Schizophr. Res.*, **7**, 225–231.
8. Gur, R.E., Petty, R.G., Turetsky, B.I. and Gur, R.C. (1996) Schizophrenia throughout life: sex differences in severity and profile of symptoms. *Schizophr. Res.*, **21**, 1–12.
9. Hafner, H. (2003) Gender differences in schizophrenia. *Psychoneuroendocrinology*, **28** (Suppl. 2), 17–54.
10. Freeman, M.P., Smith, K.W., Freeman, S.A., McElroy, S.L., Kmetz, G.E., Wright, R. and Keck, P.E., Jr (2002) The impact of reproductive events on the course of bipolar disorder in women. *J. Clin. Psychiatry*, **63**, 284–287.
11. Gattaz, W.F., Vogel, P., Riecher-Rossler, A. and Soddu, G. (1994) Influence of the menstrual cycle phase on the therapeutic response in schizophrenia. *Biol. Psychiatry*, **36**, 137–139.
12. Hallonquist, J.D., Seeman, M.V., Lang, M. and Rector, N.A. (1993) Variation in symptom severity over the menstrual cycle of schizophrenics. *Biol. Psychiatry*, **33**, 207–209.
13. Hendrick, V., Altshuler, L.L. and Burt, V.K. (1996) Course of psychiatric disorders across the menstrual cycle. *Harv. Rev. Psychiatry*, **4**, 200–207.
14. Kendell, R.E., Chalmers, J.C. and Platz, C. (1987) Epidemiology of puerperal psychoses. *Br. J. Psychiatry*, **150**, 662–673.
15. Riecher-Rossler, A., Hafner, H., Stumbaum, M., Maurer, K. and Schmidt, R. (1994) Can estradiol modulate schizophrenic symptomatology? *Schizophr. Bull.*, **20**, 203–214.
16. Seeman, M.V. (1996) The role of estrogen in schizophrenia. *J. Psychiatry Neurosci.*, **21**, 123–127.
17. Seeman, M.V. and Lang, M. (1990) The role of estrogens in schizophrenia gender differences. *Schizophr. Bull.*, **16**, 185–194.
18. Usall, J., Busquets, E., Araya, S., Ochoa, S. and Gost, A. (2000) Gender differences in schizophrenia. A literature review. *Actas Esp. Psiquiatr.*, **28**, 178–185.
19. Amin, Z., Canli, T. and Epperson, C.N. (2005) Effect of estrogen–serotonin interactions on mood and cognition. *Behav. Cogn. Neurosci. Rev.*, **4**, 43–58.
20. Berman, K.F., Schmidt, P.J., Rubinow, D.R., Danaceau, M.A., Van Horn, J.D., Esposito, G., Ostrem, J.L. and Weinberger, D.R. (1997) Modulation of cognition-specific cortical activity by gonadal steroids: a positron-emission tomography study in women. *Proc. Natl Acad. Sci. USA*, **94**, 8836–8841.
21. Greene, G.L., Gilna, P., Waterfield, M., Baker, A., Hort, Y. and Shine, J. (1986) Sequence and expression of human estrogen receptor complementary DNA. *Science*, **231**, 1150–1154.
22. Mosselman, S., Polman, J. and Dijkema, R. (1996) ER beta: identification and characterization of a novel human estrogen receptor. *FEBS Lett.*, **392**, 49–53.
23. Perlman, W.R., Tomaskovic-Crook, E., Montague, D.M., Webster, M.J., Rubinow, D.R., Kleinman, J.E. and Weickert, C.S. (2005) Alteration in estrogen receptor alpha mRNA levels in frontal cortex and hippocampus of patients with major mental illness. *Biol. Psychiatry*, **58**, 812–824.
24. Perlman, W.R., Webster, M.J., Kleinman, J.E. and Weickert, C.S. (2004) Reduced glucocorticoid and estrogen receptor alpha messenger ribonucleic acid levels in the amygdala of patients with major mental illness. *Biol. Psychiatry*, **56**, 844–852.
25. Kos, M., Reid, G., Denger, S. and Gannon, F. (2001) Minireview: genomic organization of the human ERalpha gene promoter region. *Mol. Endocrinol.*, **15**, 2057–2063.
26. Hirata, S., Shoda, T., Kato, J. and Hoshi, K. (2003) Isoform/variant mRNAs for sex steroid hormone receptors in humans. *Trends Endocrinol. Metab.*, **14**, 124–129.
27. Osterlund, M.K., Grandien, K., Keller, E. and Hurd, Y.L. (2000) The human brain has distinct regional expression patterns of estrogen receptor alpha mRNA isoforms derived from alternative promoters. *J. Neurochem.*, **75**, 1390–1397.
28. Bollig, A. and Miksicek, R.J. (2000) An estrogen receptor-alpha splicing variant mediates both positive and negative effects on gene transcription. *Mol. Endocrinol.*, **14**, 634–649.
29. Perlman, W.R., Matsumoto, M., Beltaifa, S., Hyde, T.M., Saunders, R.C., Webster, M.J., Rubinow, D.R., Kleinman, J.E. and Weickert, C.S. (2005) Expression of estrogen receptor alpha exon-deleted mRNA variants in the human and non-human primate frontal cortex. *Neuroscience*, **134**, 81–95.
30. Nielsen, D.M., Ehm, M.G. and Weir, B.S. (1998) Detecting marker-disease association by testing for Hardy–Weinberg disequilibrium at a marker locus. *Am. J. Hum. Genet.*, **63**, 1531–1540.
31. Vana, A.C., Lucchinetti, C.F., Le, T.Q. and Armstrong, R.C. (2007) Myelin transcription factor 1 (Myt1) expression in demyelinated lesions of rodent and human CNS. *Glia*, **55**, 687–697.
32. Wunderle, V.M., Critcher, R., Ashworth, A. and Goodfellow, P.N. (1996) Cloning and characterization of SOX5, a new member of the human SOX gene family. *Genomics*, **36**, 354–358.
33. Armstrong, R.C., Migneault, A., Shegog, M.L., Kim, J.G., Hudson, L.D. and Hessler, R.B. (1997) High-grade human brain tumors exhibit increased expression of myelin transcription factor 1 (MYT1), a zinc finger DNA-binding protein. *J. Neuropathol. Exp. Neurol.*, **56**, 772–781.
34. Lai, T., Jabaudon, D., Molyneaux, B.J., Azim, E., Arlotta, P., Menezes, J.R. and Macklis, J.D. (2008) SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron*, **57**, 232–247.
35. Kim, M.Y., Jeong, B.C., Lee, J.H., Kee, H.J., Kook, H., Kim, N.S., Kim, Y.H., Kim, J.K., Ahn, K.Y. and Kim, K.K. (2006) A repressor complex, AP4 transcription factor and geminin, negatively regulates expression of target genes in nonneuronal cells. *Proc. Natl Acad. Sci. USA*, **103**, 13074–13079.
36. Ohtaka-Maruyama, C., Miwa, A., Kawano, H., Kasai, M. and Okado, H. (2007) Spatial and temporal expression of RP58, a novel zinc finger transcriptional repressor, in mouse brain. *J. Comp. Neurol.*, **502**, 1098–1108.
37. Yamada, K., Gerber, D.J., Iwayama, Y., Ohnishi, T., Ohba, H., Toyota, T., Aruga, J., Minabe, Y., Tonegawa, S. and Yoshikawa, T. (2007) Genetic analysis of the calcineurin pathway identifies members of the EGR gene family, specifically EGR3, as potential susceptibility candidates in schizophrenia. *Proc. Natl Acad. Sci. USA*, **104**, 2815–2820.
38. Cheng, M.C., Liao, D.L., Hsiung, C.A., Chen, C.Y., Liao, Y.C. and Chen, C.H. (2008) Chronic treatment with aripiprazole induces differential gene expression in the rat frontal cortex. *Int. J. Neuropsychopharmacol.*, **11**, 207–216.
39. Maruyama, H., Toji, H., Harrington, C.R., Sasaki, K., Izumi, Y., Ohnuma, T., Arai, H., Yasuda, M., Tanaka, C., Emson, P.C. *et al.* (2000) Lack of an association of estrogen receptor alpha gene polymorphisms and transcriptional activity with Alzheimer disease. *Arch. Neurol.*, **57**, 236–240.
40. Aoki, K., Meng, G., Suzuki, K., Takashi, T., Kameoka, Y., Nakahara, K., Ishida, R. and Kasai, M. (1998) RP58 associates with condensed chromatin and mediates a sequence-specific transcriptional repression. *J. Biol. Chem.*, **273**, 26698–26704.
41. Fuks, F., Burgers, W.A., Godin, N., Kasai, M. and Kouzarides, T. (2001) Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J.*, **20**, 2536–2544.
42. van Duijnhoven, F.J., Bezemer, I.D., Peeters, P.H., Roest, M., Uitterlinden, A.G., Grobbee, D.E. and van Gils, C.H. (2005) Polymorphisms in the estrogen receptor alpha gene and mammographic density. *Cancer Epidemiol. Biomarkers Prev.*, **14**, 2655–2660.
43. Albagha, O.M., Pettersson, U., Stewart, A., McGuigan, F.E., MacDonald, H.M., Reid, D.M. and Ralston, S.H. (2005) Association of oestrogen receptor alpha gene polymorphisms with postmenopausal bone loss, bone mass, and quantitative ultrasound properties of bone. *J. Med. Genet.*, **42**, 240–246.
44. Ferrari, S.L. and Rizzoli, R. (2005) Gene variants for osteoporosis and their pleiotropic effects in aging. *Mol. Aspects Med.*, **26**, 145–167.
45. Shearman, A.M., Cooper, J.A., Kotwinski, P.J., Humphries, S.E., Mendelsohn, M.E., Housman, D.E. and Miller, G.J. (2005) Estrogen receptor alpha gene variation and the risk of stroke. *Stroke*, **36**, 2281–2282.
46. Fox, C.S., Yang, Q., Cupples, L.A., Guo, C.Y., Atwood, L.D., Murabito, J.M., Levy, D., Mendelsohn, M.E., Housman, D.E. and Shearman, A.M. (2005) Sex-specific association between estrogen receptor-alpha gene

variation and measures of adiposity: the Framingham Heart Study. *J. Clin. Endocrinol. Metab.*, **90**, 6257–6262.
47. Lehtimaki, T., Kunnas, T.A., Mattila, K.M., Perola, M., Penttila, A., Koivula, T. and Karhunen, P.J. (2002) Coronary artery wall atherosclerosis in relation to the estrogen receptor 1 gene polymorphism: an autopsy study. *J. Mol. Med.*, **80**, 176–180.
48. Mattila, K.M., Axelman, K., Rinne, J.O., Blomberg, M., Lehtimaki, T., Laippala, P., Roytta, M., Viitanen, M., Wahlund, L., Winblad, B. *et al.* (2000) Interaction between estrogen receptor 1 and the epsilon4 allele of apolipoprotein E increases the risk of familial Alzheimer's disease in women. *Neurosci. Lett.*, **282**, 45–48.
49. Yaffe, K., Lui, L.Y., Grady, D., Stone, K. and Morin, P. (2002) Estrogen receptor 1 polymorphisms and risk of cognitive impairment in older women. *Biol. Psychiatry*, **51**, 677–682.
50. Kohn, Y. and Lerer, B. (2005) Excitement and confusion on chromosome 6q: the challenges of neuropsychiatric genetics in microcosm. *Mol. Psychiatry*, **10**, 1062–1073.
51. Lindholm, E., Ekholm, B., Shaw, S., Jalonen, P., Johansson, G., Pettersson, U., Sherrington, R., Adolfsson, R. and Jazin, E. (2001) A schizophrenia-susceptibility locus at 6q25, in one of the world's largest reported pedigrees. *Am. J. Hum. Genet.*, **69**, 96–105.
52. Feng, J., Yan, J., Michaud, S., Craddock, N., Jones, I.R., Cook, E.H., Jr, Goldman, D., Heston, L.L., Peltonen, L., Delisi, L.E. *et al.* (2001) Scanning of estrogen receptor alpha (ERalpha) and thyroid hormone receptor alpha (TRalpha) genes in patients with psychiatric diseases: four missense mutations identified in ERalpha gene. *Am. J. Med. Genet.*, **105**, 369–374.
53. Jones, I., Middle, F., McCandless, F., Coyle, N., Robertson, E., Brockington, I., Lendon, C. and Craddock, N. (2000) Molecular genetic studies of bipolar disorder and puerperal psychosis at two polymorphisms in the estrogen receptor alpha gene (ESR 1). *Am. J. Med. Genet.*, **96**, 850–853.
54. Lai, I.C., Liao, D.L., Bai, Y.M., Lin, C.C., Yu, S.C., Chen, J.Y. and Wang, Y.C. (2002) Association study of the estrogen receptor polymorphisms with tardive dyskinesia in schizophrenia. *Neuropsychobiology*, **46**, 173–175.
55. Ouyang, W.C., Wang, Y.C., Hong, C.J. and Tsai, S.J. (2001) Estrogen receptor alpha gene polymorphism in schizophrenia: frequency, age at onset, symptomatology and prognosis. *Psychiatr. Genet.*, **11**, 95–98.
56. Tsai, S.J., Wang, Y.C., Hong, C.J. and Chiu, H.J. (2003) Association study of oestrogen receptor alpha gene polymorphism and suicidal behaviours in major depressive disorder. *Psychiatr. Genet.*, **13**, 19–22.
57. Garcia Pedrero, J.M., Zuazua, P., Martinez-Campa, C., Lazo, P.S. and Ramos, S. (2003) The naturally occurring variant of estrogen receptor (ER) ERDeltaE7 suppresses estrogen-dependent transcriptional activation by both wild-type ERalpha and ERbeta. *Endocrinology*, **144**, 2967–2976.
58. Morris, J.A., Jordan, C.L. and Breedlove, S.M. (2004) Sexual differentiation of the vertebrate nervous system. *Nat. Neurosci.*, **7**, 1034–1039.
59. Hao, J., Rapp, P.R., Leffler, A.E., Leffler, S.R., Janssen, W.G., Lou, W., McKay, H., Roberts, J.A., Wearne, S.L., Hof, P.R. *et al.* (2006) Estrogen alters spine number and morphology in prefrontal cortex of aged female rhesus monkeys. *J. Neurosci.*, **26**, 2571–2578.
60. Tang, Y., Janssen, W.G., Hao, J., Roberts, J.A., McKay, H., Lasley, B., Allen, P.B., Greengard, P., Rapp, P.R., Kordower, J.H. *et al.* (2004) Estrogen replacement increases spinophilin-immunoreactive spine number in the prefrontal cortex of female rhesus monkeys. *Cereb. Cortex*, **14**, 215–223.
61. Wang, J., Cheng, C.M., Zhou, J., Smith, A., Weickert, C.S., Perlman, W.R., Becker, K.G., Powell, D. and Bondy, C.A. (2004) Estradiol alters transcription factor gene expression in primate prefrontal cortex. *J. Neurosci. Res.*, **76**, 306–314.
62. Kulkarni, J., Riedel, A., de Castella, A.R., Fitzgerald, P.B., Rolfe, T.J., Taffe, J. and Burger, H. (2002) A clinical trial of adjunctive oestrogen treatment in women with schizophrenia. *Arch. Womens Ment. Health*, **5**, 99–104.
63. Kulkarni, J., Riedel, A., de Castella, A.R., Fitzgerald, P.B., Rolfe, T.J., Taffe, J. and Burger, H. (2001) Estrogen—a potential treatment for schizophrenia. *Schizophr. Res.*, **48**, 137–144.
64. Kos, M., Denger, S., Reid, G., Korach, K.S. and Gannon, F. (2002) Down but not out? A novel protein isoform of the estrogen receptor alpha is expressed in the estrogen receptor alpha knockout mouse. *J. Mol. Endocrinol.*, **29**, 281–286.
65. Flouriot, G., Griffin, C., Kenealy, M., Sonntag-Buck, V. and Gannon, F. (1998) Differentially expressed messenger RNA isoforms of the human estrogen receptor-alpha gene are generated by alternative splicing and promoter usage. *Mol. Endocrinol.*, **12**, 1939–1954.
66. Keks, N.A., Hill, C., Opeskin, K., Copolov, D. and Dean, B. (1999) Psychiatric diagnosis after death: the problems of accurate diagnosis from case history review and relative interviews. Dean, B. and Kleinman, J.E. (eds), *Using CNS Tissue in Psychiatric Research: A Practical Guide.* Overseas Publishers Association, Amsterdam, pp. 19–37.
67. Zalcman, S., Endicott, J., Clayton, P. and Winokur, G. (1983) *Diagnostic Evaluation After Death (DEAD).* National Institute of Mental Health, Rockville.
68. Lipska, B.K., Deep-Soboslay, A., Weickert, C.S., Hyde, T.M., Martin, C.E., Herman, M.M. and Kleinman, J.E. (2006) Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. *Biol. Psychiatry*, **60**, 650–658.
69. Torrey, E.F., Webster, M., Knable, M., Johnston, N. and Yolken, R.H. (2000) The stanley foundation brain collection and neuropathology consortium. *Schizophr. Res.*, **44**, 151–155.
70. Weickert, C.S., Hyde, T.M., Lipska, B.K., Herman, M.M., Weinberger, D.R. and Kleinman, J.E. (2003) Reduced brain-derived neurotrophic factor in prefrontal cortex of patients with schizophrenia. *Mol. Psychiatry*, **8**, 592–610.
71. Kaufmann, C.A., Suarez, B., Malaspina, D., Pepple, J., Svrakic, D., Markel, P.D., Meyer, J., Zambuto, C.T., Schmitt, K., Matise, T.C. *et al.* (1998) NIMH genetics initiative millenium schizophrenia consortium: linkage analysis of African-American pedigrees. *Am. J. Med. Genet.*, **81**, 282–289.
72. Kozlovsky, N., Shanon-Weickert, C., Tomaskovic-Crook, E., Kleinman, J.E., Belmaker, R.H. and Agam, G. (2004) Reduced GSK-3beta mRNA levels in postmortem dorsolateral prefrontal cortex of schizophrenic patients. *J. Neural. Transm.*, **111**, 1583–1592.
73. Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M. and Werner, T. (2005) MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, **21**, 2933–2942.
74. Mathias, R.A., Gao, P., Goldstein, J.L., Wilson, A.F., Pugh, E.W., Furbert-Harris, P., Dunston, G.M., Malveaux, F.J., Togias, A., Barnes, K.C. *et al.* (2006) A graphical assessment of *P*-values from sliding window haplotype tests of association to identify asthma susceptibility loci on chromosome 11q. *BMC Genet.*, **7**, 38.
75. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.