Chapter 5

Genotyping Two Candidate CNVs in an Extended Case-Control Cohort

5.1 Candidate I: *CHL1* (Close Homolog of L1) at 3p26

5.1.1 Functional Significance of *CHL1*

The first CNV genotyped in the extended replication cohort was located 5' to the gene Close homolog of L1 on chromosome 3p. Close homolog of L1 (*CHL1* or *CALL*) belongs to the L1 family of cell adhesion molecules, a subgroup of the immunoglobulin superfamily (Leshchyns'ka et al. 2006). The mammalian L1 family comprises four neural recognition molecules: *L1*, *CHL1*, neurofascin and *NrCAM*. Together with other families of neural recognition molecules, such as cadherins and integrins, the four L1 family proteins play important roles in neural development, neuronal regeneration and synaptic activities (Maness and Schachner 2007). Members of the L1 family have been implicated in a wide spectrum of brain abnormalities and neuropsychiatric disorders. For instance, *L1CAM* is responsible for several X-linked mental retardation syndromes, as well as increased risk of schizophrenia (Kurumaji et al. 2001). *NrCAM* is the binding partner of the synapse associated protein *SAP102 (Davey et al. 2005)*, a MAGUK family member involved in non-syndromic X-linked mental retardation (Laumonnier et al. 2007). Moreover, allelic variants of *NrCAM* have been implicated in autism (Hutcheson et al. 2004; Sakurai et al. 2006; Marui et al. 2008) and addiction vulnerability (Ishiguro et al. 2006).

CHL1 is expressed in subpopulations of neurons and glial cells in both the central and peripheral nervous system during early development, starting from embryonic day 13 in mice (Hillenbrand et al. 1999). The protein is involved in a variety of neuronal functions including: (i) the promotion of neurite outgrowth (Holm et al. 1996; Hillenbrand et al. 1999), (ii) neuronal migration (Buhusi et al. 2003; Demyanenko et al. 2004; Nishimune et al. 2005) and survival (Chen et al. 1999); (iii) modulation of

synaptic plasticity (Montag-Sallaz et al. 2002; Morellini et al. 2007); (iv) synaptic targeting via the clathrin dependent pathway (Leshchyns'ka et al. 2006); and (v) the regulation of axon regeneration (Chaisuksunt et al. 2000; Zhang et al. 2000; Jakovcevski et al. 2007).

CHL1 was demonstrated to have involvement in mental retardation and schizophrenia in both human and mouse genetic studies. In human, *CHL1* was one of the genes deleted in the 3p-syndrome characterized by mental and psychomotor retardation with congenital anomalies. Haplo-insufficiency of the gene was suggested to be responsible for the cognitive phenotype in this syndrome (Angeloni et al. 1999). *CHL1* was also identified at the translocation breakpoints in a patient with non-specific mental retardation (Frints et al. 2003). In addition, two association studies have established a link between a mis-sense polymorphism in the signal peptide region of *CHL1* (Leu17Phe) and schizophrenia (Sakurai et al. 2002; Chen et al. 2005).

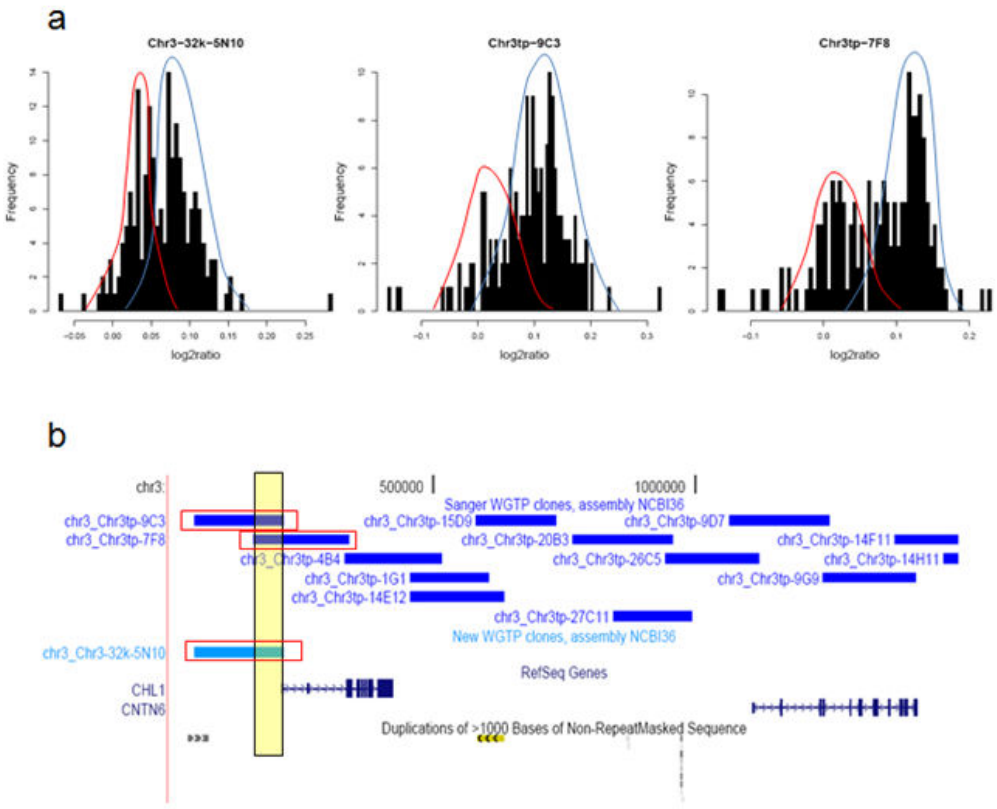
In *CHL1* knockout mice, a delayed reactivity to environmental stimuli was observed, suggesting a deficit in extracting relevant information from the environment in these mice (Morellini et al. 2007). Chl1-/- mice were also less aggressive and have altered social behavior in exploration tests including open field, resident-intruder test and urine marking (Frints et al. 2003; Morellini et al. 2007). In sensorimotor gating functions, Chl1-/- mice exhibited impaired prepulse inhibition (PPI) of the acoustic startle response (Irintchev et al. 2004), a key endophenotype of schizophrenia. In terms of cognitive ability, synaptic plasticity in juvenile Chl1-/- mice showed enhanced perisomatic inhibitory synaptic transmission, coupled with impaired long-term potentiation (LTP) in the hippocampus (Nikonenko et al. 2006). In adults, hippocampus synaptic plasticity was normal, but there was an enhanced basal synaptic transmission in the dentate gyrus (Morellini et al. 2007). Finally, in the Morris

Water Maze test, Chl1-/- mice displayed an altered exploratory pattern (Montag-Sallaz et al. 2002; Frints et al. 2003), suggesting cognitive processes involving spatial memory could be impaired. Altogether, *CHL1* knockout mouse models suggested mutations of the gene could result in schizophrenia by disrupting key neurological functions related to disease phenotypes.

5.1.2 Structure of the *CHL1* 5' Deletion Polymorphism

A copy number polymorphism 5' of *CHL1* was detected via the log2 ratio distributions of three consecutive clones at the most distal end of chromosome 3p (Chr3-32K-5N10, Chr3tp-9C3, & Chr3tp-7F8) (Figure 5.1). Visually at least two peaks could be detected in the log2ratio histograms of the three affected clones (Figure 5.1a & b). Furthermore, the log2ratio distributions for three clones showed strong positive correlation with each other (Pearson correlation of log2ratio >0.6), but not with other neighbouring clones (Figure 5.1c).

By plotting the log2ratio of one of the three clones against another and applying bivariate clustering algorithm, the data clustered into three classes (Figure 5.2). Cluster A and B were non-discrete, with only a small shift in log2ratio, indicate the presence of a small CNV. We hypothesized that the 3 clusters represented normal (cluster A), heterozygous deleted (cluster B) and homozygous deleted (cluster C) samples respectively, with the reference DNA sample (HapMap NA10851) being heterozygous deleted.



c

Clone	Position on Chr3	Pearson Correlation with Clone Chr3tp-9C3
Chr3-32k-5N10	132207.5	0.62811
Chr3tp-9C3	132217.0	1.00000
Chr3tp-7F8	251366.0	0.61511
Chr3tp-4B4	425974.0	0.09403
Chr3tp-1G1	533672.5	-0.08985
Chr3tp-14E12	547467.0	-0.00283
Chr3tp-15D9	658724.5	0.02715
Chr3tp-20B3	863307.0	0.05340
Chr3tp-27C11	920543.5	0.11604
Chr3tp-26C5	1031880.0	-0.01170

Figure 5.1 A copy number polymorphism at chromosome 3p26.3 was detected by three BAC clones (a) Log2ratio histograms* of the 3 BAC clones at 3p26.3: *Chr3-32k-5N10* (left), *Chr3tp-9C3* (middle) and *Chr3tp-7F8* (right). Bi-allelic distribution (two peaks) in the histograms indicates a common CNV in the genomic region. (**b)** Location of the three clones (red box) and their overlapping region (yellow box) with respect to the location of the gene *CHL1*. **(c)** Pearson correlation coefficients of log2ratio indicated high correlation of *Chr3tp-9C3* with *Chr3-32k-5N10* and *Chr3tp-7F8,* but not with other neighbouring clones.

* histograms generated from combined SCZ & LBC cohort data

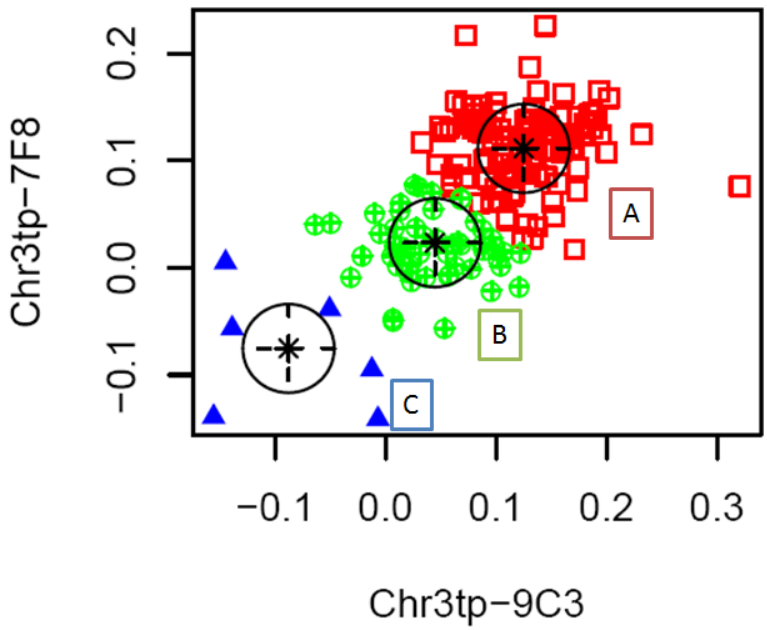


Figure 5.2 CHL1 CNV genotypes from bivariate clustering. Applying a model-based hierarchical clustering algorithm on clones Chr3tp-7F8 and Chr3tp-9C3 resulted in three major clusters, which correspond to (A) normal, (B) heterozygous deletion and (C) homozygous deletion genotypes.

The location of the *CHL1* CNV was determined and breakpoints were sequenced in five HapMap samples that were either heterozygous or homozygous carriers of the deletion polymorphism (NA12144, NA12892, NA18576, NA18976 and NA10851) (Figure 5.3). All samples generated the same breakpoint structure, with distal and proximal breakpoints localised to *Alu* repeat sequences (Figure 5.3a). The CNV occurred at approximately 4000 bp 5' of the gene *CHL1*, with the distal breakpoint estimated at chr3:204,000, and the proximal breakpoint at chr3:209,500. Recently released high-resolution data from the Nimblegen oligonucleotide array platform (Genome Structural Variation Consortium 2008) further confirmed the deletion size in the heterozygous reference DNA sample NA10851 (Figure 5.3b). The size of the CNV was estimated to be around 5.5 kb.

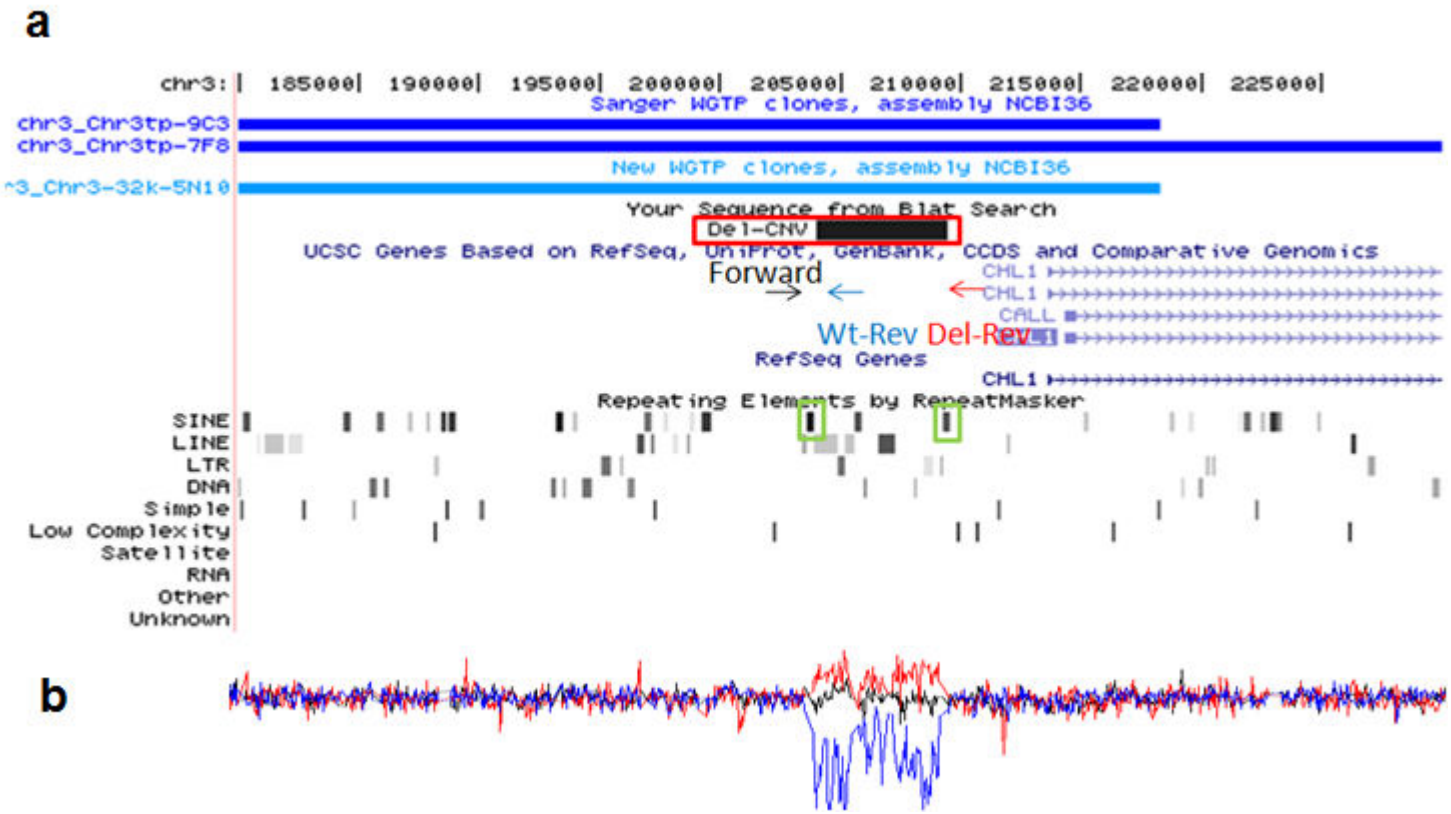


Figure 5.3 Genomic location of the *CHL*1 5' CNV. The CHL1 5' deletion polymorphism (black box) was validated using both PCR and Nimblegen oligonucleotide array. **(a)** PCR determined the CNV breakpoints to fall on a pair of SINE *Alu* elements (green boxes). Arrows represent the locations of the PCR primer series used for CNV genotyping and validation. →forward primers; ← reverse primers for the wild-type allele; ← reverse primers for the deletion allele **(b)** Nimblegen oligonucleotide array confirmed the size and location of CNV to be chr3:204,000-209500 and revealed both deletions and duplications in different samples.

5.1.3 Correlation of *CHL1* CNV to mRNA Expression

To assess the relationship of this 5' deletion polymorphism with *CHL1* mRNA expression, we extracted gene expression data from a recent publication (Stranger et al. 2007) assessing the effect of genetic variation (SNPs and CNVs) on expression phenotypes for 210 unrelated individuals from International HapMap Project (IHMC 2005).

In the analysis performed by Stranger *et al.* on a whole genome expression array platform (Illumina Sentrix Human-6 Expression BeadChip), the expression of the *CHL1* probe *GI_27894375-S* was shown to correlate with the presence of a CNV within the *CHL1* gene (at clone *Chr3tp-4B4*). This correlation was one of the five most significant associations between CNV and expression phenotypes in the CEU population as detected at the 0.05 permutation threshold. *Chr3tp-4B4* covers the 3' end of CHL1, and is immediately downstream of the three clones (*Chr3-32K-5N10*, *Chr3tp-9C3*, & *Chr3tp-7F8*) that span the 5.5 kb deletion polymorphism we described.

We next investigated expression data of the clones that span the deletion polymorphism of interest (*Chr3tp-9C3, & Chr3tp-7F8)*[3], by comparing them to published HapMap WGTP array CGH data from the same samples (Redon et al. 2006). As shown in Figure 5.4, correlation between array CGH log2ratio and the expression level of the *CHL1* probe (*GI_27894375-S)* was most prominent in CEU samples (COR=0.519, Pearson Correlation). Correlation was less significant in YRI or CHB +JBP samples.

[3] the third clone *Chr3-32K-5N10* was not present in the HapMap array CGH data and therefore was not analyzed

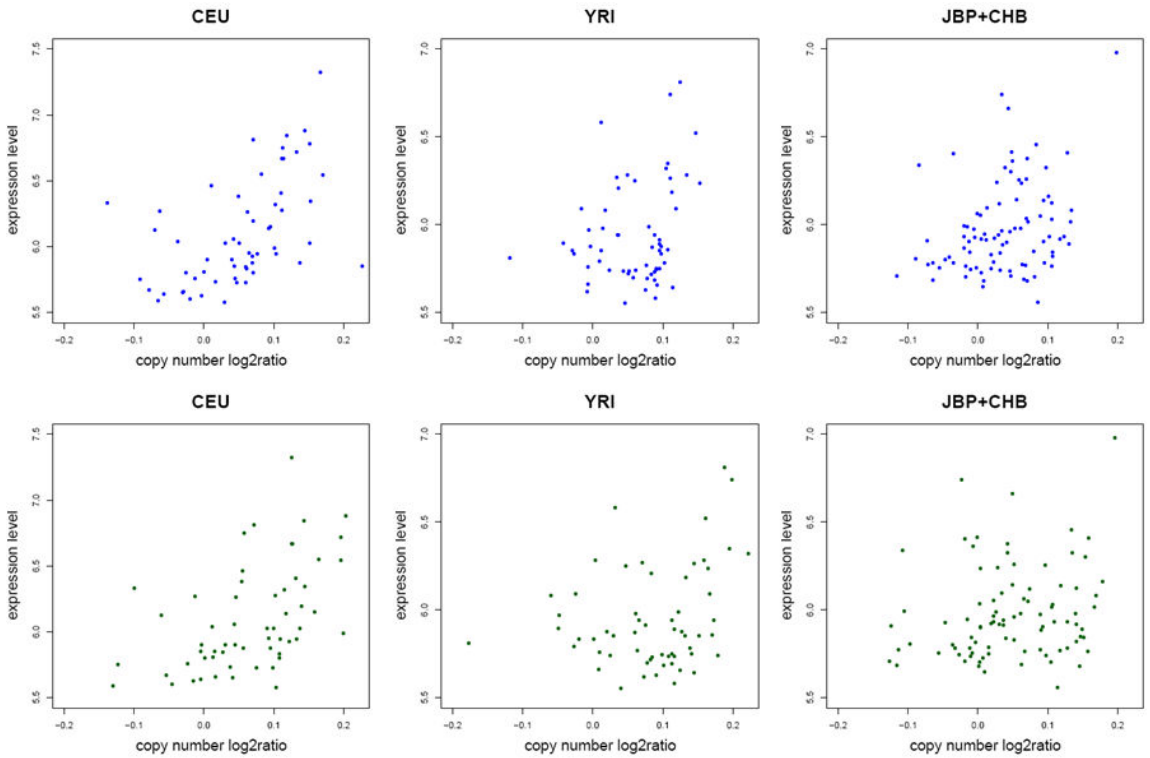


Figure 5.4 *CHL1* expression level against BAC clone log2ratio in HapMap samples. *CHL1* expression level (as measured by probe *GI_27894375-S* on Illumina Sentrix Human-6 Expression BeadChiparray) was plotted against log2ratio data from three BAC clones spanning the *CHL1* 5' CNV. Expression data was plotted against log2ratio of clones: **(a)** *Chr3tp-9C3 &* **(b)** *Chr3tp-7F8*

Left: CEU- Utah samples with European ancestry; Middle: YRI- Yoruba samples with African ancestry; Right: JBP+CHB Japanese and Chinese samples with Asian ancestry

All HapMap CEU samples were then genotyped for the *CHL1* 5' deletion polymorphism using a PCR strategy with one forward primer and two reverse primers spanning the deletion, generating a ~700 bp PCR band for a normal allele, or a ~400 bp band for the deletion allele. PCR genotyping allowed each sample to be categorized into discrete CNV genotypes of: homozygous deletion (0), heterozygous deletion (1) or normal (2) (Figure 5.5).

Considering only the CEU samples with European ancestry (Figure 5.6), the correlation of expression with clone log2ratio was 0.519 (Pearson Correlation) (Figure 5.6a). The comparison of expression with discrete CNV genotypes also showed a positive association with a p-value of 0.0436 (Figure 5.6b). In summary, our analysis revealed a potential positive correlation between a novel 5' *CHL1* CNV with expression of the *CHL1* gene in individuals of European ancestry.

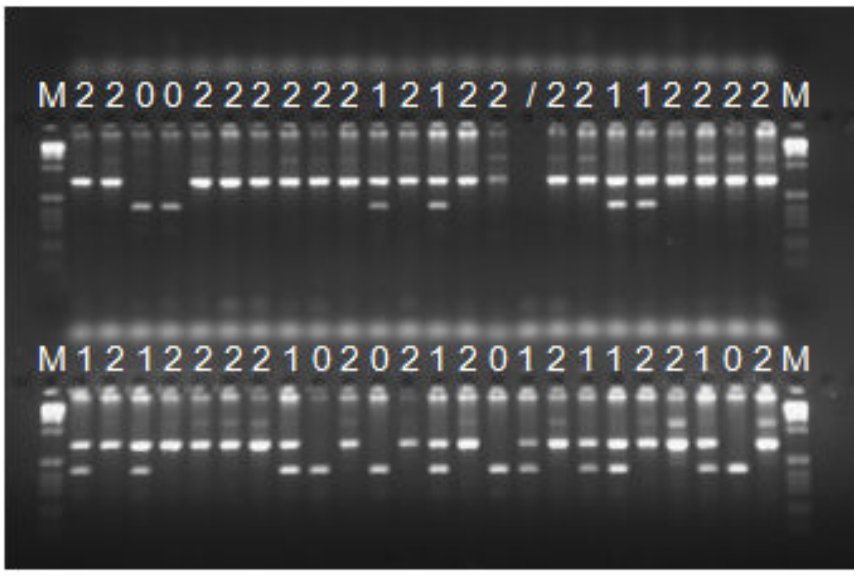


Figure 5.5 Genotyping of the *CHL1* 5' deletion polymorphism in HapMap samples. Genotyping was performed by PCR on 90 HapMap CEU samples. The ~700 bp band represents the "normal" allele and the ~400 bp band represents the deletion allele.

(M: marker; 0: homozygous deletion; 1: heterzygous deltion; 2: normal allele)

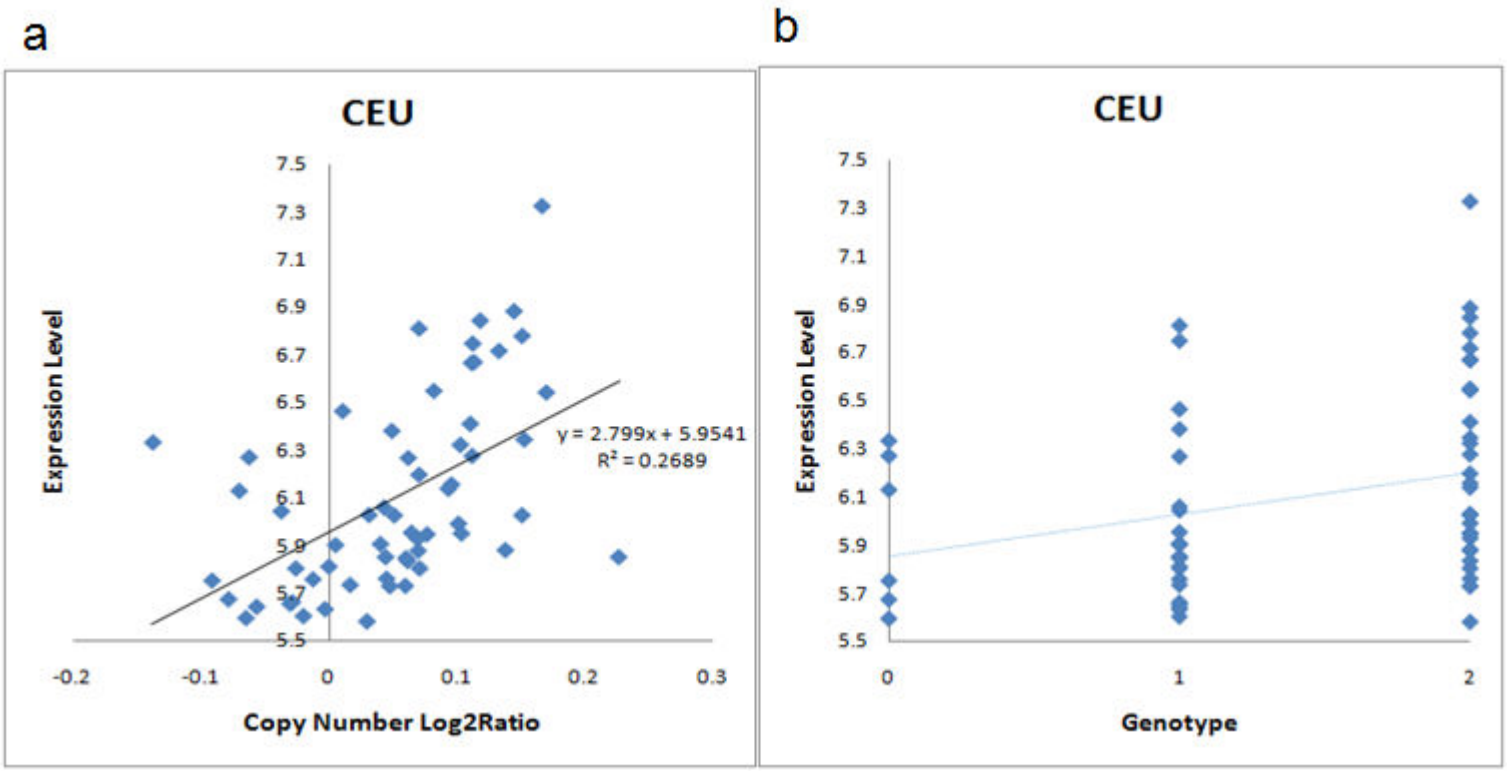


Figure 5.6 *CHL1* gene expression against copy number of the *CHL1* 5' CNV in HapMap CEU samples. (a) *CHL1* expression plotted against array CGH log2ratio signal demonstrate positive association (Pearson correlation= 0.519). **b)** *CHL1* expression plotted against discrete CNV genotypes resulted from PCR demonstrates positive association (p-value= 0.0436)

(0: homozygous deletion; 1: heterozygous deletion; 2: normal)

5.1.4 Genotyping CHL1 CNV in a Larger Case-Control Cohort

Our initial whole-genome schizophrenia CNV screen on 91 cases and 92 controls revealed possible association of the deletion polymorphism with disease status (p-value=0.00876) (section 4.6 & Table 5.1). Prompted by this enrichment of *CHL1* 5' deletion polymorphism in patients, and the correlation of *CHL1* expression with the 5' CNV genotypes, we attempted to replicate such findings in an extended case-control cohort.

354 SCZ DNA and 376 LBC DNA samples were genotyped using the same PCR strategy as described above. CNV genotype distribution [-/-: +/-: +/+] in the disease cohort was [12 (*3%*): 125 (*33%*): 239 (*64%*)], compared with [11 (*3%*): 121 (*34%*): 222 (*63%*)] in the control cohort (Table 5.2). The result did not show any association between CNV genotypes with disease (chi-square test, p-value = 0.964). Allele frequency distribution also demonstrated a lack of association (chi-square test, p-value = 0.906, data not shown). In conclusion, despite the initial positive association and the relevant biological evidence, we were not able to extend the findings of *CHL1* 5' deletion enrichment in disease to a larger case-control cohort.

Table 5.1 *CHL1* 5' CNV distribution determined by the original genome-wide array CGH data. CNV genotypes for 91 schizophrenia patients and 92 controls were estimated by log2ratio bivariate clustering of clones *Chr3tp-9C3* and *Chr3tp-7F8*. Samples with high uncertainty* (uncertainty value > 0.1), mostly between clusters A and B, were removed from analysis. A significant association of deletion carriers to disease status was observed (p-value = 0.00876)

CNV Status	SCZ	LBC
Deletion Carrier (Cluster B + C)	34 (45.9%)	19 (24.4%)
No Deletion (Cluster A)	40 (54.1%)	59 (75.6%)
Total	74	78

**uncertainty values (0-1) were generated for each data point by the model-based hierarchical clustering algorithm in R Mclust package*

.

Table 5.2 *CHL1* 5' CNV distribution in an extended case-control cohort determined by PCR genotyping. (a) Results for 376 DNA samples from the Schizophrenia Cohort. **(b)** Results for 354 DNA samples from the Lothian Birth Control Cohort. No association of the CNV with disease was detected (p-value=0.984).

a

SCZ Cohort	Plate DN30417	Plate DN30418	Plate DN30419	Plate DN30420	SUMMARY CNV Status	No. of samples	% of Samples
Homozygous Deletion	5	2	3	2	Homozygous Deletion	12	3.19%
Heterozygous Deletion	30	33	26	36	Heterozygous Deletion	125	33.24%
Normal (No Deletion)	59	59	64	57	Normal (No Deletion)	239	63.56%
Total	**94**	**94**	**93**	**95**		**376**	**100%**

b

LBC Cohort	Plate DN30396	Plate DN30398	Plate DN30399	Plate DN30400	SUMMARY CNV Status	No. of samples	% of Samples
Homozygous Deletion	1	5	4	1	Homozygous Deletion	11	3.11%
Heterozygous Deletion	31	35	20	35	Heterozygous Deletion	121	34.18%
Normal (No Deletion)	56	51	64	51	Normal (No Deletion)	222	62.71%
Total	**88**	**91**	**88**	**87**		**354**	**100%**

5.2 Candidate II: *CHRFAM7A* (CHRNA7-FAM7A fusion gene) at 15q13

The second candidate genotyped in the replication cohort was a deletion polymorphism at *CHRFAM7A*. The fusion gene *CHRNA7-FAM7A* (*CHRFAM7A*) and its counterpart alpha7 neuronal nicotinic acetylcholine receptor gene (*CHRNA7*) reside at the 15q13-14 locus, a region that has been widely studied in the genetics of schizophrenia. Linkage analysis and association studies have implicated this region in schizophrenia and its endophenotypes (see Table 5.4), and in a number of neurological conditions including bipolar disorder (Edenberg et al. 1997), epilepsy (Elmslie et al. 1997) and Alzheimer's Disease (Carson et al. 2008).

5.2.1 Genomic Architecture at chromosome 15q13-14

The 15q13-14 locus has a complex genomic structure with segmental duplications and an evolutionary history of structural rearrangements (Riley et al. 2002; Sharp et al. 2008) (Figure 5.7). *CHRNA7* and *CHRFAM7A* are ~1.6 Mb apart in this locus. *CHRFAM7A* is part of the segmental duplication BP4 (breakpoint 4 of the Prader-Willli Syndrome), which has three large regions showing high homology (≥95%) with BP5 (breakpoint 5) (Figure 5.7 panel 4), the downstream segmental duplication immediately adjacent to *CHRNA7* (Sharp et al. 2008). *CHRFAM7A* was partially derived from *CHRNA7* and partially from *FAM7A*, a novel gene which was known to have duplicated to several locations on chromosome 15 (Riley et al. 2002). The last 5 exons of *CHRNA7* and *CHRFAM7A* share >99% homology (Gault et al. 1998). The formation of this hybrid gene *CHRFAM7A* was suggested to be a recent event specific to the human evolutionary history (Riley et al. 2002; Taske et al. 2002), since *CHRFAM7A* is absent in other primates such as the chimpanzee (Locke et al. 2003).

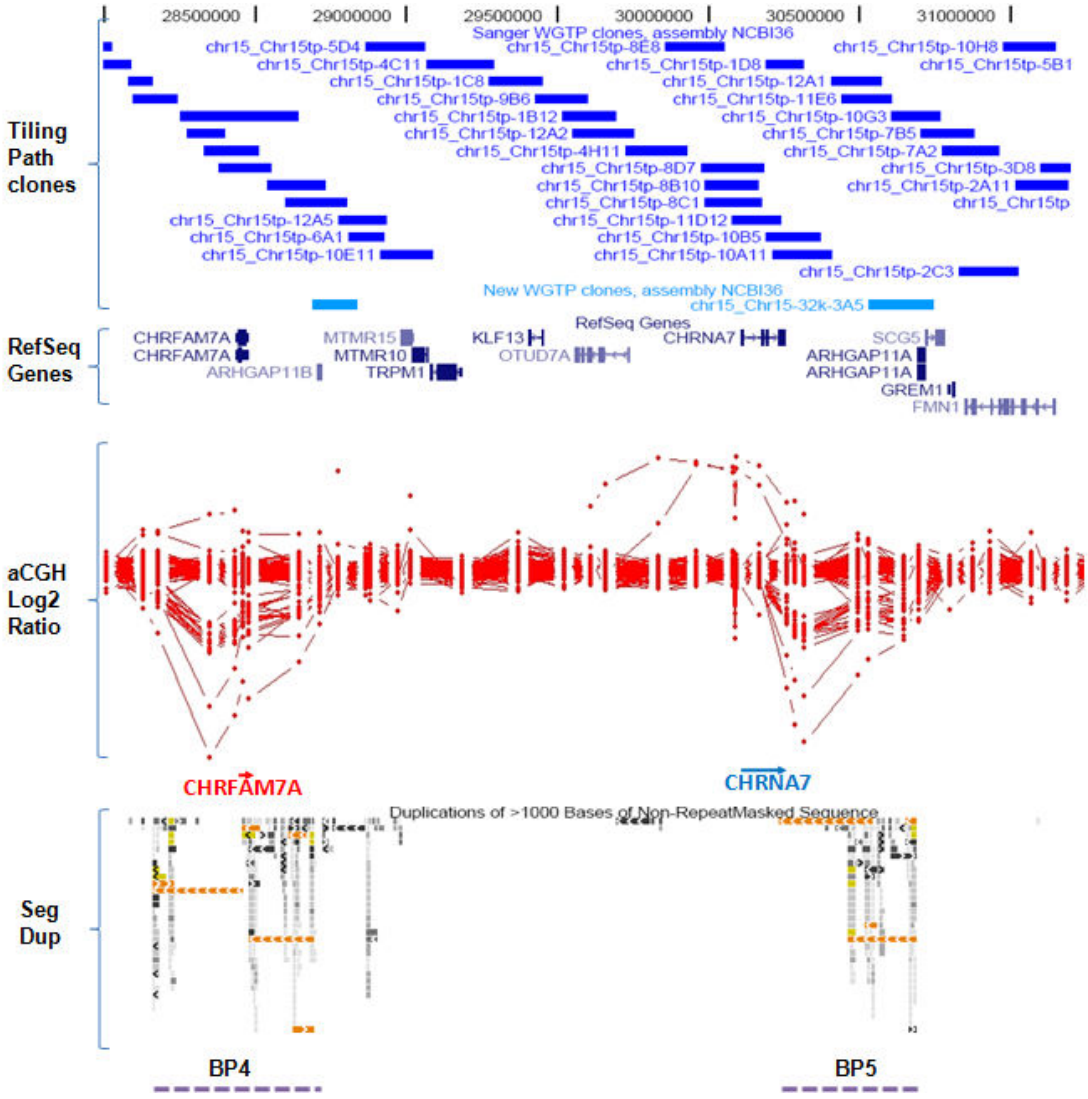


Figure 5.7 The 15q13-14 genomic locus. *(from top to bottom)* WGTP array BAC clones (blue blocks) (panel 1); RefSeq genes (panel 2); Array CGH data for 91 schizophrenia patients (Each red line represents the genomic profile of one patient) (panel 3); segmental duplication structures aligned to genomic locus (panel 4). The location of BP4 and BP5 (breakpoint 4 and breakpoint 5 of the Prader Willi Syndrome) are shown as purple dotted lines. They represent a pair of segmental duplication blocks with high homology to each other.

CHRNA7 and *CHRFAM7A* are shown in opposite orientations in the public DNA sequence database (NCBI Build 36), suggesting an inversion during the emergence of *CHRFAM7A* in the human lineage. Flomen *et al.* presented a model of the evolution of *CHRFAM7A* as a series of deletions, inversions and duplications, with postulated ancestral and intermediate structures (Flomen et al. 2008). It was also proposed that the *CHRFAM7A* exists as polymorphic inversion structures in the current population (consistent with previous findings from Riley et. al), and that the *CHRFAM7A* null allele represents the persisting ancestral chromosome. The 3 known allele variants of *CHRFAM7A* in humans are depicted in Figure 5.8.

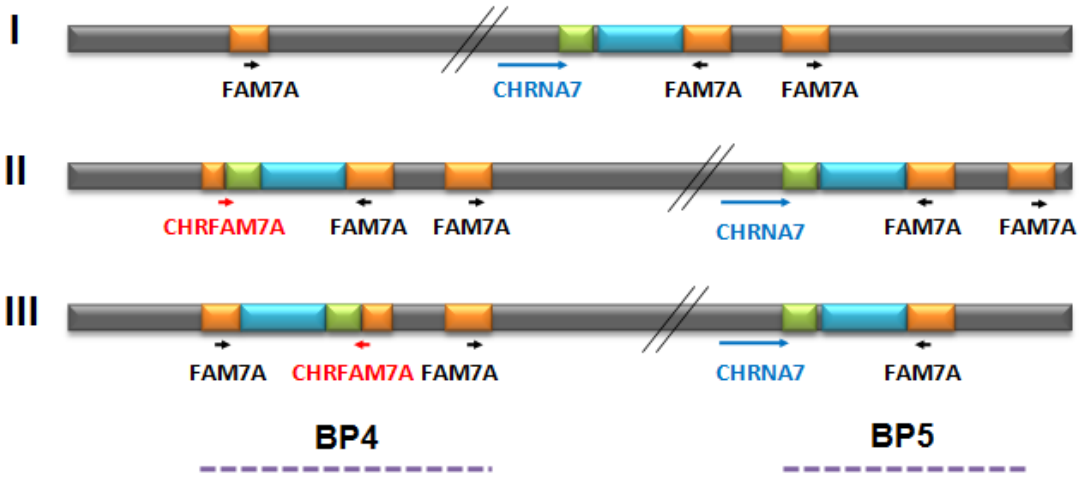


Figure 5.8 The three existing polymorphic structures of *CHRFAM7A* allele at 15q13-14. Structures as proposed by Flomen *et al.:* The first structure (I) is a null allele with no *CHRFAM7A*, possibly representing the ancestral allele which has persisted in the current population. The second structure (II) was formed subsequent to the acquisition of the hybrid gene *CHRFAM7A*. *CHRFAM7A* and *CHRNA7* are in the same orientation. The third structure (III) represents the inverted allele with respect to structure II, with *CHRNA7* and *CHRFAM7A* in opposite orientation. This is the allele as shown in the current genome build (NCBI36).

Structural rearrangements at 15q13-14 have been linked to several clinical phenotypes. Sharp *et al.* identified the 1.5 Mb recurrent deletion from BP4 to BP5, together with an overlapping 3Mb deletion, in patients with mental retardation and epilepsy (Sharp et al. 2008). Two recent genome-wide investigations of copy number variations in schizophrenia also identified the 15q13.3 recurrent deletions as a susceptibility locus (see section 1.6) (ISC 2008; Stefansson et al. 2008). With regard to the polymorphic *CHRFAM7A* gene, copy number of this gene was investigated in a combined cohort of patients (n=453) diagnosed with schizophrenia, bipolar affective disorder, schizoaffective or other psychoses (Flomen et al. 2006). When all disorders were considered as a single psychotic phenotype, the null *CHRFAM7A* allele was significantly associated with disease (P = 0.04) (association was not significant when the psychoses were analyzed separately). Furthermore, a 2-bp deletion in exon 6 of *CHRNA7*, which has recently been demonstrated as a surrogate marker of the inversion polymorphism of the upstream *CHRFAM7A* gene (Flomen et al. 2008), has previously been associated with bipolar disorder and major depressive disorder (Lai et al. 2001a; Lai et al. 2001b; Hong et al. 2004). In two other independent association analyses, the same 2 bp-deletion was involved in P50 sensory gating (Raux et al. 2002) and episodic memory performance (Dempster et al. 2006), both endophenotypes of schizophrenia, although association was not detected in another study with schizophrenia and other psychoses (Flomen et al. 2006). Combined evidence from these studies indicates that structural polymorphisms at chromosome 15q13-14 are likely candidates for a number of psychiatric disorders.

5.2.2 Molecular Genetic Studies linking 15q13-14 to Schizophrenia

In 1997, Freedman *et al.* first reported genetic linkage of the locus to a psychiatric phenotype (Freedman et al. 1997). In the study with nine multiplex families, a microsatellite marker in intron 2 of *CHRNA7* (D15S1360) was found to co-segregate with auditory evoked potential deficit (P50), an endophenotype related to sensory gating frequently observed in schizophrenia patients. The linkage signal to P50 was strong (LOD score of 5.9), but the signal was much weaker (LOD score decrease to 1.33) when the linkage was analyzed with the disease itself. The same microsatellite marker was then genotyped in other schizophrenia patient cohorts and the positive linkage was replicated by some (Leonard et al. 1998; Riley et al. 2000; Liu et al. 2001) but not in all studies (see Table 5.3).

Subsequently a number of linkage and association studies were performed (Table 5.3), some demonstrating positive linkage or association of the region (and of *CHRNA7* or *CHRFAM7A* polymorphisms) to schizophrenia, periodic catatonia, impaired auditory gating deficit, and/or other psychiatric disorders. However, inconsistent linkage results and lack of associations were also reported. In particular, in a recent large-scale association studies by Sanders *et. al* (1,870 cases and 2002 controls) (Sanders et al. 2008), as well as another comprehensive association study with linkage disequilibrium analysis (Iwata et al. 2007), both groups specifically investigated the *CHRNA7* gene/region, and found no association of any SNP or haplotype block with schizophrenia.

In summary, reports on genetic linkage analysis and association studies in the past decade have largely demonstrated positive results. Together with the recent compelling evidence of genomic rearrangement of this region in schizophrenia patients, the 15q13-14 locus is strongly suggested as a plausible candidate for schizophrenia pathogenesis.

Nevertheless, the specific mutation or polymorphism functionally involved in disease etiology and pathway (as opposed to markers or genomic regions) has yet to be identified.

Table 5.3 Previous reports associating the 15q13-14 locus to schizophrenia and related psychiatric disorders.

Author	Year	Study Type	Genes Under Investigation	Disease/ Phenotype	Sample Size	Result
Freedman	1997	Linkage	CHRNA7	SCZ (with P50 sensory gating deficit)	9 multiplex families	Microsatellite Marker D15S1360 in intron2 of CHRNA7 cosegregates with P50 sesory gating deficity with a LOD= 5.3. A smaller LOD score (Z=1.3) with disease
Leonard	1998	Linkage	CHRNA7	SCZ	20 families	Significant linkage at marker D15S1360 (P < 0.0024)
Neves-Pereira	1998	Linkage	CHRNA7	SCZ, SCZ AFF and others	5 multiplex families	No evidence of linkage Did not replicate findings previous postive finding with Did not replicate findings with D15S1360
Curtis	1999	Linkage	CHRNA7	SCZ	54 families	No evidence of linkage with 3 markers D15S1360, L76630, and ACTC Did not replicate previous positive findings
Riley	2000	Linkage	CHRNA7	SCZ	20 families	Microsatellite markers spanning 15q were genotyped Positive nonparametric linkage signals dectected at D15S1360,D15S1010 and D15S1043 Significant transmission disequilibrium detected at D15S1360
Stober	2000	Linkage	CHRNA7	SCZ with periodic catatonia	12 multiplex families	Evidence of linkage at the chromosome 15q candidate region (LOD score 2.75-2.89)
Stassen	2000	Association	CHRNA7	SCZ, BP and SCZ AFF	127 cases, 129 controls	Confirmed association between marker D15S1360 and schizophrenia. Also significant association with marker L76630 (P<0.0001). No significant difference between smokers and non-smokers
Gejman	2001	Linkage	CHRNA7	SCZ	68 families	20 microsatellite markers spanning chromosome 15 detected linkage at 15q11.2-q25. Single markers D15S659, ACTC, D15S150 and D15S1039 were implicated
Liu	2001	Linkage	CHRNA7	SCZ	52 families	6 microstallete markers at 15q11-14 genotyped. Positive linkage detected at D15S976 and D15S1360
Freedman	2001	Linkage	CHRNA7	SCZ (with P50 sensory gating deficit)	61 parent-child triads in 50 families	Significant linkage disequilibrium (P < 0.007) found at D15S165 within 1Mb of CHRNA7A and CHRFAM7A
Xu	2001	Linkage	CHRNA7	SCZ	31 families or trios	3 microsatellite markers genoypted. D15S165 and L76630 shows evidence of linkage
Tsuang	2001	Linkage	CHRNA7	SCZ, SCZ AFF	392 cases	Linkage detected between D15S165 and D15S1010
Meyer	2002	Linkage	CHRNA7	SCZ with periodic catatonia	1 multiplex family	Evidence strongly support linkage at 15q13-22 in a multiplex family, but gene CHRNA7 was excluded
Gault	2003	Mutation Screening & Association	Region surrounding CHRNA7 & CHRFAM	SCZ	171 cases (families), 185 control	33 polymorphisms were found in patients and controls. 21 variants in exons, but non-synonymous changes were rare
Iwata	2007	LD Analysis & Association Study	CHRNA7 and CHRFAM7A	SCZ	188 cases (188 controls)	14 microsatellite markers geneotyped, one deviate from HWE in patients but not in control. Overall no linkage detected

Shi J.	2007	Family Based Association	19 cholinergic genes	BP and SCZ AFF (BP type)	I. 474 cases (152 families) II. 83 cases (22 multiplex families)	Some SNPs showed nominally significant transmission bias, but none reach genome-wide significance after correction by multiple testing
Sanders	2008	Association	CHRNA7 (and 14 other SCZ genes)	SCZ	1,870 cases (2002 controls)	No significant association with any of the 14 genes using tag SNPs and SNPs previously reported to be associated with SCZ
Association Study of the Promoter Region of CHRNA7						
Leonard	2002	Mutation Screening & Association	Promoter Region of CHRNA7	SCZ	166 cases and 165 controls	Evidence of linkage of functional promoter variants with SCZ ((trasncription analysis subsequent to genoytping)
Houy	2004	Mutation Screening & Association	Promoter Region of CHRNA7	P50 sensory gating deficit	111 cases, 85 controls	Significant association of the -194 C allele (T/C ratio <0.45) with P50 but not SCZ. Did not replicate previous result of core promoter polymorphism from Leonard et. al.
Li	2004	Mutation Screening & Association	Promoter Region of CHRNA7	SCZ	249 cases, 273 controls	New polymorphisms were identified at the promoter region but no association detected
Association Study of the 2-bp deleiton in exon 6 of CHRNA7						
Lai	2001	Association	CHRNA7 and CHRFAM7A	Major Depressive Disorder	72 cases, 103 controls	Mildly significant association of the 2-bp deletion genotypes (P=0.027) and alleles (P=0.037) with major depressive disorder
Lai	2001	Association	CHRNA7 and CHRFAM7A	SCZ	146 cases, 151 controls	No significant association of 2-bp deletion with disease
Hong	2004	Association	CHRNA7 and CHRFAM7A	BP	77 cases, 135 controls	Mildly significant association of 2-bp deletion with BP (P=0.044)
Raux	2002	Association	CHRNA7 and CHRFAM7A	SCZ & P50 sensory gating deficit	70 cases, 77 controls	2bp deletion is not associated with SCZ, but is associated with the P50 gating deficit
Flomen	2006	Association	CHRNA7 and CHRFAM7A	SCZ, BP and SCZ AFF	453 cases, 197 controls	Mildly significant association (P = 0.04) of CHRFAM deletion genotype with disease, not significant after multiple testing. No association of the 2-bp deletion at exon 6 of CHRN7A w ith disease
Dempster	2006	Association	CHRNA7 and CHRFAM7A	SCZ with Episodic Memory Performance	251 cases	significant association of the 2bp deletion with episodic memory performance in SCZ patients
Association Study of Nicotinic Receptor Genes with Smoking in Schizophrenia						
De Luca	2004	Association	CHRNA7	SCZ with smoking	177 patients (108 smokers, 69 non smokers)	Marker D15S1360 associated with smoking in schizophrenia patients (p=0.015)
Faraone	2004	Association	CHRNA7 and 15 nicotinic acetylcholine receptor genes	SCZ with smoking		Significant linkage of the group of genes to SCZ (p = 0.039). CHRN7A alone shows marginally significant association (p = 0.095)

5.2.3 Functional Significance of *CHRNA7* and *CHRFAM7A*

As with other nicotinic acetylcholine receptors (nAChRs), *CHRNA7* is a ligand-gated ion channel that binds nicotine with low affinity. The receptor admits calcium ions into cells and mediates fast signal transmission at synapses. nACHRs are thought to assemble with other subunits to form a pentameric holoreceptor (Lindstrom 1996), although *CHRNA7* was demonstrated to form homo-oligomeric channel receptors in vitro (Couturier et al. 1990; Leonard and Bertrand 2001). *CHRNA7* plays important roles in development and in synaptic plasticity (Dani and Bertrand 2007).

Recent biochemical studies have linked the alpha7 receptor to pathways and molecules with high relevance to schizophrenia. Hancock *et al.* established the role of neuregulin-Erb signalling in targeting *CHRNA7* for axonal surface expression (Hancock et al. 2008). In a study of the CA1 region of hippocampus, the brain-derived neurotrophic factor (*BDNF*), another well-accepted schizophrenia candidate, was shown to decrease the nicotinic receptor's response in interneurons (Fernandes et al. 2008). Furthermore, antisense knockdown of *CHRNA7* in the rat demonstrated deficits in learning and memory (Curzon et al. 2006), while overexpression of the receptor in mice was suggested to improve spatial memory-related performance (Ren et al. 2007). Impaired attention (Young et al. 2007) and sensory gating functions (Young et al. 2007) were also reported in mouse models.

The function of the hybrid gene *CHRFAM7A* is not entirely understood. The gene is transcribed and mRNA detected in various tissues including the brain (Gault et al. 1998), although *CHRFAM7A* is expressed at a lower level in the brain compared to that of *CHRNA7* (Agulhon et al. 1999; De Luca et al. 2006). De Luca *et al.* also compared postmortem expression of the *CHRNA7* and *CHRFAM7*A in schizophrenia, bipolar and

controls and found a decreased *CHRNA7*:*CHRFAM7A* mRNA ratio in bipolar disorder (but not in schizophrenia) compared to controls (De Luca et al. 2006). It had been proposed that *CHRFAM7A* could affect the oligomerization of alpha-7 receptor channel, or have other gain of function or dominant negative effect on *CHRNA7* (Raux et al. 2002).

5.2.4 Copy Number Polymorphism at the *CHRFAM7A* Region

A copy number polymorphism spanning the gene *CHRFAM7A* was detected via the log2 ratio distributions of a series of clones at 15q13.2-13.3. The copy number region covers the location of the segmental duplication blocks breakpoint 4 (BP4), which has high homology to the more distal breakpoint BP5 (refer to section 5.2.1).

Applying bivariate-clustering algorithm to *Chr15tp-2G5* and *Chr15tp-2G2,* the two BAC clones that span the gene *CHRFAM7A*, two major clusters were observed (Figure 5.9). Since previous literature indicates the presence of a null *CHRFAM7A* allele in the current population (Flomen et al. 2006), we postulated that the upper cluster (cluster A) represents samples with two copies of the "normal" allele each having one copy of *CHRFAM7A*, while the lower cluster (cluster B) represents samples carrying deletion of the *CHRFAM7A* allele (the "ancestral" allele).

We assigned each of the 91 SCZ DNA samples and 92 LBC DNA samples from the array CGH screen to one of the two clusters (Figure 5.10). The ratio of samples in cluster A: cluster B in the schizophrenia cohort was 70:19 (normal:deletion) (n=89, with two samples in between clusters removed due to high uncertainty), compared to the ratio of 82:10 (normal:deletion) in the control. The enrichment of cluster B samples in the schizophrenia cohort indicates an excess of null (deletion) *CHRFAM7A* allele in the

disease population (p-value = 0.054682). We therefore genotyped the deletion polymorphism in an extended cohort of 300+ patients with equivalent number of matched controls.

With Taqman qPCR primers and probes spanning the breakpoint betjween *FAM7A* exons (D' to A) and the *CHRNA7* exons (5-10), we were able to genotype the copy number of the hybrid gene *CHRFAM7A* independent of the copy number of *CHRNA7*. The primers and probes were first tested on Hapmap individuals with known log2 ratio, and in each sample qPCR fold change correlated well with fold change estimated from array CGH log2 ratio (data not shown).

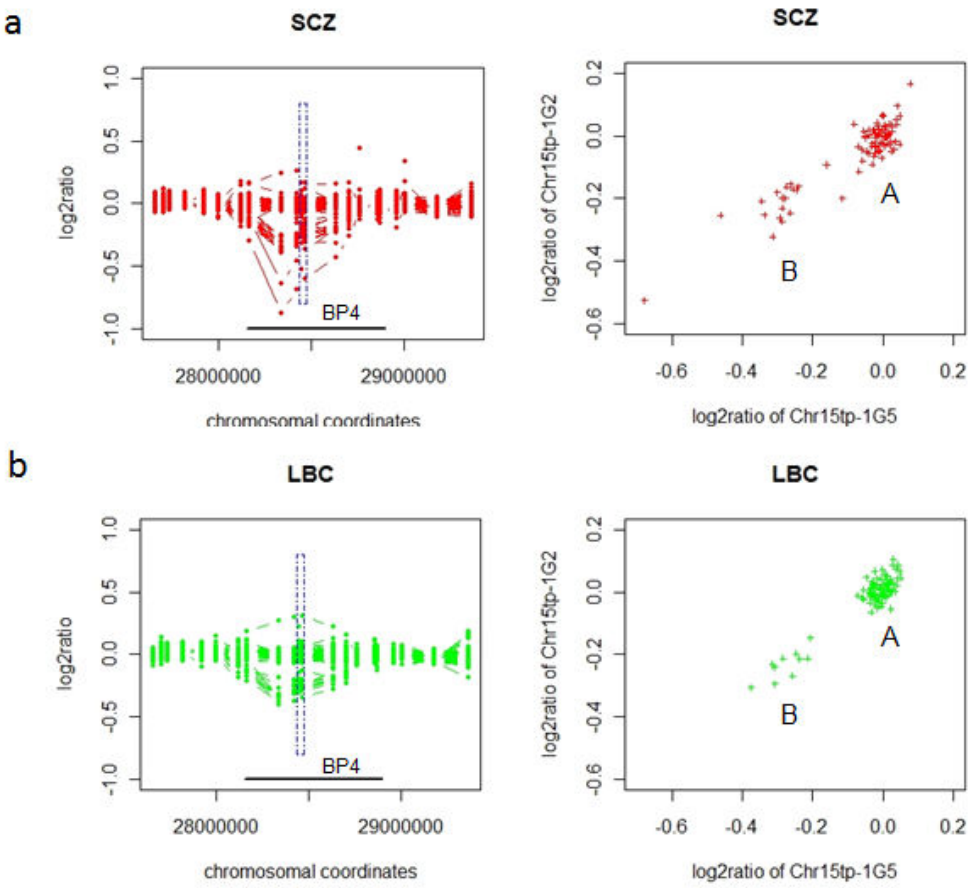
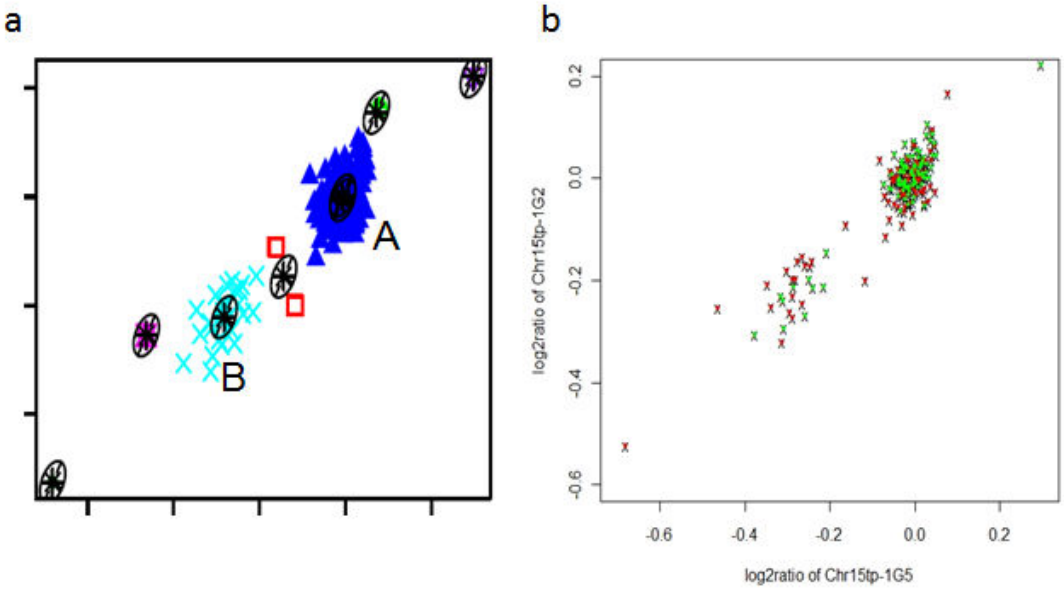


Figure 5.9 Copy number polymorphism at 15q13.2-13.3 spanning the gene *CHRFAM7A*. Array CGH data for **a)** SCZ cohort (n=91) & **b)** LBC cohort (n=92). On the left panel log2ratio was plotted against the genomic position. Location of *CHRFAM7A* was highlighted with the blue box. BP4 (breakpoint 4 of the Prader-Willi Syndrome) was shown as black line. On the right panel log2ratio of two clones spanning *CHRFAM7A* (Chr15tp-1G2 and Chr15tp-1G5) were plotted against each other. Two major clusters were detected in both SCZ and LBC cohorts.



c

	Cluster A ("Normal")	Cluster B ("Deletion")	Total
SCZ	70	19	89
LBC	82	10	92
	Chisq.test	P-value	0.054682

Figure 5.10 *CHRFAM7A* CNV genotype distributions in cases versus controls. (a) CNV genotypes were estimated from array CGH log2ratio by bivariate clustering of BAC clones *Chr15tp-1G2* and *Chr15tp-1G5*. **b)** Assigning SCZ (red) and LBC (green) samples to the two major genotype clusters. **c)** Sample counts and chi-square test statistics* show a mildly significant association of deletion carriers with disease (p=0.05).

**uncertainty values (0-1) were generated for each data point by the model-based hierarchical clustering algorithm in R Mclust package. Two samples with high uncertainty values (between cluster A and B) were removed before statistical analysis.*

Next we genotyped the *CHRFAM7A* CNV in an extended replication cohorts using qPCR. Results were obtained for 262 schizophrenia patients and 270 matched controls (each with 2 replications). A typical qPCR experiment was shown in Figure 5.11 (see Appendix D-4 for the complete set of qPCR results). Table 5.4 summarizes the genotype frequency distributions in cases and controls.

Our genotyping attempt in an extended cohort suggested that CNV genotypes at *CHRFAM7A* have no statistically significant association with disease (by genotype frequency: chi-square test, p-value = 0.693; by allele frequency: chi-square test, p-value=0.672), although by direct observation there seem to be a slight trend of increased deletion allele in the disease cohort.

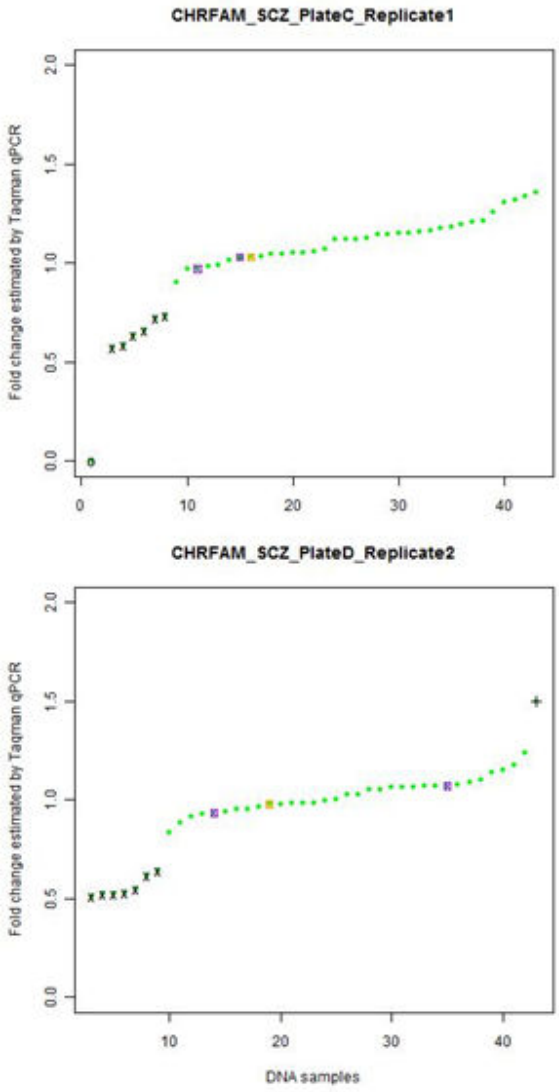


Figure 5.11 Analysis of a typical Taqman qPCR experiment to determine *CHRFAM7A* genotypes. In each 96-well plate 40 samples (+2 controls) were genotyped. Fold change of all samples in the same experiment (sorted in ascending order) were plotted on the same graph (green dots) and CNV genotypes were estimated by the change of the slope of the graph.

o: Homozygous deletion samples; x: heterozygous deletion samples; +: duplication samples : control sample NA10851; : control sample NA12776

Table 5.4 *CHRFAM7A* CNV genotype distributions in the extended case and control cohorts. qPCR genotyping results for **(a)** SCZ replication cohort (n=270) (**b)** LBC replication cohort (n=262). Chi-square test of the CNV genotype distributions suggests no statistical significant difference between the two cohorts (p-value=0.0693).

a

	Plate	Homodel	Hetdel	Normal	Duplication	Total
1	SCZ_PlateA	0	7	32	1	40
2	SCZ_PlateB	0	6	32	0	38
3	SCZ_PlateC	1	6	30	0	37
4	SCZ_PlateD	0	7	29	1	37
5	SCZ_PlateE	3	4	31	2	40
6	SCZ_PlateF	0	7	32	0	39
7	SCZ_PlateH	1	4	34	4	39
	No. of Samples	**5**	**41**	**220**	**8**	**270**
	% of Samples	**1.85%**	**15.19%**	**81.48%**	**2.96%**	

b

	Plate	Homodel	Hetdel	Normal	Duplication	Total
1	LBC_PlateA	0	4	34	1	39
2	LBC_PlateB	0	7	22	1	30
3	LBC_PlateC	0	8	29	0	37
4	LBC_PlateE	0	5	32	2	39
5	LBC_PlateF	2	6	30	1	39
6	LBC_PlateG	0	4	33	1	38
7	LBC_PlateH	0	4	36	0	40
	No. of Samples	**2**	**38**	**216**	**6**	**262**
	% of Samples	**0.76%**	**14.50%**	**82.44%**	**2.29%**	

5.3 Chapter Summary and Discussion

This chapter describes targeted genotyping and characterization of two schizophrenia candidate copy number polymorphisms, and an attempt to replicate findings from our initial array CGH screen in an extended case-control cohort.

The *CHL1* 5' deletion polymorphism at 3p26 was fully characterized by a combination of CGH data clustering and PCR strategies, together with published gene expression and high-resolution CNV data. We demonstrated that the small (~5 kb), frequent deletion located 5' upstream of *CHL1* could partially explain expression level of the gene. The deletion was significantly associated with schizophrenia in our initial screen on 91 patients and 92 controls, but the enrichment was not observed in the extended cohort of 300+ patients and matched controls.

The second candidate, a copy number polymorphism at *CHRFAM7A* at 15q13, demonstrated recurrent deletion and duplication in the population. The highly complex genomic architecture with segmental duplication and inversion made it difficult to devise a simple PCR assay for CNV genotyping. A Taqman qPCR strategy with specific primers and probes for *CHRFAM7A* was used to determine gene dosage. Our initial whole-genome CNV screen demonstrated an enrichment of the deletion phenotypes. In an extended replication cohort, qPCR genotypes for *CHRFAM7A* were obtained for 250+ cases and matched controls. No significant association was detected between CNV genotypes and disease, although there was slight bias of deletion genotypes by direct observation.

Neither CNV polymorphism showed significant association in the extended cohort of ~300 cases and ~300 controls. We postulate the following to explain our inability to replicate the initial associations:

First, our initial CNV screen was modest in sample size, and could result in statistical bias leading to false-positive associations. It is worth mentioning that we see a high plate-to-plate variation (irrespective of disease status) in our subsequent PCR genotyping of the *CHL1* 5' CNV (see Table 5.2), indicating the relevance of random population sampling. This emphasizes the need for replication in CNV association studies, for instance a two-stage approach of polymorphism identification with subsequent replication, as in most SNP association studies. Replication is particular important for studies with small initial sample size (e.g. for cost-effective genome-wide CNV discoveries as in our study), and is pertinent to detecting associations of frequent copy number polymorphisms, as well as the relatively rare but recurrent copy number changes.

Alternatively, the common polymorphism candidates we targeted could have real impact on disease risks but with small effect sizes. The lack of association in subsequent replication attempts may point to the need for yet larger samples to statistically detect such association with a small odds ratio. This relates to the common-disease common-variant (CD-CV) hypothesis (Collins et al. 1997), which suggests that common disease traits like schizophrenia could be the consequence of common allelic variants with small to modest effects, typically under less stringent selection pressure (Carlson et al. 2004; Zondervan and Cardon 2004). Consistent with such an hypothesis, in the case of *CHRFAM7A* CNV, we see a non-statistically significant bias of patient samples towards

deletion genotypes, which could represent a small increased disease risk as conferred by the deletion allele.

Finally, we cannot rule out non-sampling-related artifacts in the original CNV screen and analysis. For example, in the initial screen, accuracy of copy number quantification is compromised by converting continuous measurements (fluorescent intensity-based data) into discrete CNV genotypes (McCarroll and Altshuler 2007). In addition, variability in CNV measurements between cohorts, e.g. due to DNA quality, as implicated in our study, could have a potential impact on genotype classification (Barnes et al. 2008). Knowing the true underlying genotype is important to validate association signals (Todd 2006). To this end, one could use targeted quantification methods such as Taqman qPCR to deliver a more acceptable accuracy level in larger-scale genotyping, a strategy we employed. Recently, more robust, high-throughput and low-cost methods have been developed and proved valuable in disease association studies, even for multi-allelic loci with a wide range of copy numbers (see section 1.2 for discussion of such methods). Improved statistical methods have also been applied to intensity-based data for more accurate CNV genotype classifications (Barnes et al. 2008).