We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

5,000 Open access books available
124,000 International authors and editors
140M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Impact of Epidemiology on Molecular Genetics of Schizophrenia

Nagafumi Doi, Yoko Hoshi, Masanari Itokawa, Takeo Yoshikawa and Tomoe Ichikawa

1. Introduction

Schizophrenia is a common deleterious psychosis that begins typically in late adolescence or early adulthood (Gottesman, 1991; Jablensky, 1995). Although it has a strong genetic component in its etiology, no susceptibility genes conferring a large proportion of heritability have been identified (Allen et al., 2008; Need et al., 2009). Results of association studies including genome-wide scans have been inconsistent, and schizophrenia-associated genes including copy number variations differ across populations or even across individuals of the same ethnicity (Allen et al., 2008; Xu et al., 2008; Need et al., 2009). Thus, the situation of molecular genetics of schizophrenia has become rather perplexing just contrary to our expectation.

In this chapter, we describe the consistent major epidemiological findings of schizophrenia, and show how these evident macroscopic aspects shed light to the confused microscopic aspects of schizophrenia genetics today, proposing a new hypothesis for this puzzling disorder.

2. Devil’s triangle of human genetics – Epidemiological facts of schizophrenia

We describe here three epidemiological facts of schizophrenia – high prevalence, high heritability and low reproductive fitness. These properties form a Devil’s triangle; any combination of the two tends to exclude the third, and in this triangle most diseases vanish except for schizophrenia, suggesting that schizophrenia has a unique etiological basis among the many human diseases.

2.1 Schizophrenia as a common disease

Substantial evidence of epidemiology shows that schizophrenia crosses all cultures and tribes in different continents at a relatively high prevalence (approximately 0.7%; 95%
Confidence Interval 0.3% - 2.7%) (Saha et al., 2005); the prevalence of schizophrenia, at the macro-level, varies within narrow limits (Jablensky, 1995), and appears to be stable across generations in several countries (Harrison et al., 1991; Osby et al., 2001). This epidemiological fact suggests that schizophrenia has an ancient origin.

### 2.2 Schizophrenia as a heritable disease

It has long been known that schizophrenia runs in families (McGue & Gottesman, 1991). Adoption studies demonstrate an increased risk of schizophrenia in biological relatives of adoptees with schizophrenia, suggesting that genetic components play an important role in the etiology of schizophrenia (Kendler & Dichl, 1993). Now it has been established by twin studies that heritability of schizophrenia is ~0.85 (Cannon et al., 1998; Cardno et al., 1999).

Although the mode of transmission of schizophrenia is still unknown, several reports suggest a higher maternal transmission of schizophrenia (Shimizu et al., 1987; Goldstein et al., 1990; Valero et al., 1998; Li et al., 2007).

### 2.3 Schizophrenia as a low fitness disease

It has been well documented that the fertility of patients with schizophrenia, particularly of males, is remarkably reduced compared to healthy individuals (Böök, 1953; Larson & Nyman, 1973; Ødegård, 1980; Nanko & Moridaira, 1993; Fànanás & Bertranpetit, 1995; Nimgaonkar, 1998; McGrath et al., 1999; Haukka et al., 2003; Svensson et al., 2007). The latest meta-analysis (Bundy et al., 2011) shows that fertility ratio (patients/controls) is ~0.39 and that the reduction of fertility is more pronounced in males (male/female ratio is ~0.54).

Because schizophrenia is an early onset disease (late adolescence ~ early adulthood), psychotic symptoms of the disease such as autistic way of life and abnormal behaviors may make mating unsuccessful. This tendency may be more pronounced in males because the age at onset is significantly lower in males than in females (Jablensky, 1995; Kulkarni & Fink, 2000). Thus, unsuccessful mating, coupled with an increased mortality (McGrath et al., 2008), may remarkably reduce the fertility of patients with schizophrenia.

### 2.4 Schizophrenia and the Devil’s triangle of human genetics

The three epidemiological characteristics of schizophrenia - high prevalence, high heritability and low fitness - form a Devil’s triangle; any combination of the two tends to exclude the third, and in this triangle most diseases vanish except for schizophrenia (Fig. 1). Diseases with high prevalence and high heritability such as type 2 diabetes and adult cancers are late-onset diseases and exhibit almost normal reproductive fitness. Diseases with high heritability and low reproductive fitness such as most harmful Mendelian diseases in childhood are rare. Diseases with low reproductive fitness and high prevalence such as poor nutrition, severe injuries and infections in childhood or early adulthood are mainly due to the environmental factors.

This may lead us to strongly suspect that schizophrenia has a unique etiological basis among the many human diseases.
3. Persistence problem and mutation-selection balance in schizophrenia

The three epidemiological characteristics of schizophrenia mentioned above give a paradox. How has a highly heritable disease associated with a remarkable biological disadvantage never been extinct in the long human history? And how can it persist at a relatively high prevalence? This ‘persistence problem of schizophrenia’ (or ‘schizophrenia paradox’) has puzzled scientists for long years (Huxley et al., 1964; Crow, 1995; Brüne, 2004; Keller & Miller, 2006).

In this section, we discuss that the only plausible mechanism for the persistence is mutation-selection balance with or without heterozygote advantage. Based on the consistent epidemiological findings on the fertility of patients with schizophrenia and their family members, we show that heterozygote advantage works in the mitochondrial genome model but not in the nuclear genome model.

3.1 Mutation-selection balance is the only plausible mechanism for the persistence

From an evolutionary viewpoint, four explanations are possible for the persistence: (i) ancestral neutrality, (ii) negative frequency-dependent selection, (iii) heterozygote advantage (balancing selection or pleiotropy), and (iv) mutation-selection balance.

‘Ancestral neutrality’ assumes that reproductive fitness of affected individuals and/or their relatives was higher in ancient environments and that selection coefficients of pathogenic alleles were close to zero. Because the effective population size in ancient times might be much smaller than now, pathogenic but neutral or almost neutral alleles could be fixed by genetic drift. While this hypothesis explains that schizophrenia has not been extinct in the long human history, ancestral neutrality itself provides no explanation for the apparently stable prevalence of the disease across generations today: although ‘ancestral neutrality’ might be plausible, it needs another mechanism to account for the persistence in modern environments, where the effective population size has been expanded and the influence of negative selection pressure may be much stronger than ever before.
'Negative frequency-dependent selection' explains the persistence only when the fitness of the affected individuals increases as the prevalence in the general population decreases, which seems not to be the case with schizophrenia.

Thus, the remaining possibility for persistence mechanism is mutation-selection balance with or without heterozygote advantage.

3.2 Heterozygote advantage works in the mitochondrial genome model but not in the nuclear genome model for schizophrenia

‘Heterozygote advantage’ assumes that the susceptibility alleles increase the fitness of the unaffected gene carriers, thereby sustaining the gene frequencies. This line of explanations may include: (i) physiological advantage (resistance to shock, infections, and poor nutrition etc.), (ii) a higher sexual activity and/or attractiveness, and (iii) creative intelligence or a higher trait creativity including ‘everyday creativity’.

This hypothesis needs two lines of confirmation: (a) that unaffected gene carriers have such advantages, and (b) that such advantages really contribute to sufficiently increase their reproductive fitness.

It seems to gain the confirmation (a). For example, Erlenmeyer-Kimling (1968) reported an increased survival rate of female children of parents with schizophrenia, proposing a possible physiological advantage associated with schizophrenia. Kinney et al. (2001), in a well-designed and methodologically sophisticated study, showed that an advantage of everyday creativity was linked to a subtle clinical picture (schizotypal signs) in a non-psychotic sample of schizophrenia offspring.

However, it lacks the confirmation (b) in the nuclear genome model. This hypothesis, although theoretically plausible and fascinating, has not been supported by most epidemiological studies, which show a decreased reproductive fitness of unaffected siblings of patients with schizophrenia. Although recent large-sampled epidemiological studies (Bassett et al., 1996; McGrath et al., 1999; Haukka et al., 2003; Svensson et al., 2007) have consistently shown that the reproductive fitness of unaffected female siblings of patients with schizophrenia is slightly but significantly increased (1.02-1.08), it is not large enough to compensate for the gene loss due to the decreased reproductive fitness of patients (0.2-0.3 in males and 0.4-0.5 in females) and their unaffected male siblings (0.9-1.0) in the nuclear genome model. On the other hand, the latest meta-analysis (Bundy et al., 2011) shows no significant difference between the fertility of parents of patients with schizophrenia and healthy controls, although there is a trend towards parents having more offspring. Therefore, heterozygote advantage seems not to work in the nuclear genome model.

On the other hand, it works in the mitochondrial genome model because mitochondrial DNA (mtDNA) is transmitted to the next generation only through females. Indeed, we can see that this slightly elevated reproductive fitness of the unaffected female siblings, coupled with the less pronounced decreased reproductive fitness of female patients, is sufficient to compensate for the gene loss; when we calculate $\Delta$, the cross-generational reduction of the frequency of females with a putative pathogenic mtDNA in the general population, using the data in the largest-sampled cohort study to date (Haukka et al., 2003), we have $-\Delta < 5.06 \times 10^{-5}$ (Note). This figure implies that the gene loss can be balanced by de novo
mutation in the mtDNA which occurs at a rate of $8.8 \times 10^{-4} \sim 1.3 \times 10^{-2}$ per locus per generation ($4.3 \times 10^{-3}$ on average) (Sigurðardóttir et al., 2000).

4. Persistence criterion for nuclear susceptibility genes for schizophrenia

As is shown in the previous section, putative pathogenic genes, if located in the mtDNA, are sustained by mutation-selection balance with heterozygote advantage. On the other hand, if located in the ncDNA, they should be sustained by mutation-selection balance without heterozygote advantage. In this section, we introduce our previous work (Doi et al., 2009), in which we carefully re-examined the necessary conditions for putative nuclear susceptibility genes for schizophrenia and deduced a criterion (persistence criterion, or ‘P-criterion’) that every nuclear susceptibility gene should fulfill for persistence of the disease, and present its applications to association studies for schizophrenia.

4.1 Three basic assumptions

At first we describe our three basic assumptions.

4.1.1 An ideal human population

We assume here a random-mating human population with a sufficiently large effective population size at equilibrium, where negative selection pressures on the susceptibility alleles for schizophrenia are predominant and the effect of genetic drift is negligibly small. The prevalence $p$ of schizophrenia in this ideal human population is assumed to be stable across generations by mutation-selection balance. Therefore, the gene frequency in the general population ($m_G$) is given in terms of the gene frequencies in the affected population ($m_A$) and in the unaffected population ($m_U$):

$$m_G = pm_A + (1-p)m_U,$$

or

$$m_A - m_G = (1-p)d.$$

(1)

4.1.2 Mutation-selection balance in each risk locus

We assume here that the total of the population frequencies of the pathogenic alleles at each risk locus is preserved by mutation-selection balance. Therefore, $-\Delta m_G$, the cross-generational reduction of the frequency of a pathogenic allele should not be more than the rate of mutations that produce pathogenic variants at the locus. On the other hand, since mutations at the locus include mutations of two directions that produce pathogenic or non-pathogenic variants, the mutation rate at the locus ($\mu$) should be greater than the rate of mutations that produce pathogenic variants at the locus.

Thus we have:

$$\mu > -\Delta m_G.$$  

(2)

4.1.3 Multifactorial threshold model

We assume the multifactorial threshold model, in which quantitative traits such as liability to the disease are determined by multiple genetic and non-genetic factors including a
stochastic and/or an epigenetic effect. Under this assumption, the relative fitness as a quantitative trait in the affected population is determined by multiple factors and approximately follows a gamma distribution with a mean \(1 - s\). (\(s\) is the selection coefficient of schizophrenia; the mean relative fitness in the normal population is 1.)

Distribution curve of the reproductive fitness in the affected population

Distribution curve of the reproductive fitness in the affected subpopulation with a schizophrenia-associated allele \(M\) never shifts to the right unless \(M\) has a strong protective effect (i.e. an effect of elevating carrier’s reproductive fitness by reducing severity of and liability to the disease). Therefore, we can assume that \(s_M\), the selection coefficient in the affected subpopulation with a schizophrenia-associated allele \(M\), is not smaller than \(s\) \((s \leq s_M < 1)\) for a susceptibility allele (Fig. 2). The inequality \(s > s_M\) implies that \(M\) is a resistance gene that reduces severity and risk of the disease.

No special assumptions else are required on the allelic structure in each locus, penetrance of each susceptibility gene, and possible interactions among the loci.

4.2 Deduction of the P-criterion

Now we proceed to deduce the P-criterion. From the assumptions, \(m'_G\), the population frequency of the schizophrenia-associated allele \(M\) in the next generation, is given by:

\[
m'_G = \frac{p \cdot m_A \cdot (1 - s_A) + (1 - p) \cdot m_{G} \cdot 1}{p \cdot (1 - s_M) + (1 - p) \cdot 1} = \frac{m_G - s_M p m_A}{1 - s_M p}.
\]

Therefore the reduction of the population frequency of the schizophrenia-associated allele \(M\) per generation is:

\[
-\Delta m_G = m_G - m'_G = \frac{s_M p (m_A - m_G)}{1 - s_M p} = p(1 - p) d \cdot \frac{s_M}{1 - s_M p}.
\]
Impact of Epidemiology on Molecular Genetics of Schizophrenia

From (2) and (3) we have:

\[ \mu > p(1-p)d \frac{s_M}{1-s_mp} . \]  

(4)

Since \( \frac{s_M}{1-s_mp} \) is monotonically increasing for \( s_M (0 < s_M < 1) \) and \( s \leq s_M < 1 \) holds for the susceptibility allele \( M \), we have:

\[ \mu > p(1-p)d \frac{s}{1-sp} , \text{ or } \frac{(1-sp)\mu}{(1-p)sp} > d . \]  

(5)

On the other hand, the principle of association studies demands: \( 0 < d \).

Thus we have the criterion for a susceptibility gene:

\[ 0 < d < \nu , \text{ where } \nu = \frac{(1-sp)\mu}{(1-p)sp} . \]  

(6)

From the observation (5), we can see that \( d\geq \nu \) implies \( s > s_M \) for any schizophrenia-associated variant \( M \) which is sustained by mutation-selection balance.

4.3 Parameter estimate for schizophrenia

Mutation rates on autosomes and the X chromosome almost always fall within the range between \( 10^{-6} \) and \( 10^{-4} \) per locus per generation (usually < \( 10^{-4} \); one generation = 20 years) (Nachman & Crowell, 2000) and can be approximated by a linear function of the parental age at least under 30 years for maternal age and under 40 years for paternal age (Risch et al., 1987). Large-sampled cohort studies in Israel, Sweden and Denmark show that the mean age of parents in the general population is ~ 28 years for mothers and ~31 years for fathers; the mean age of both parents is < 29.6 years (Malaspina et al., 2001; El-Saadi et al., 2004). Therefore we assume here:

\[ 10^{-6} < \mu < \frac{29.6}{20} \times 10^{-4} = 1.48 \times 10^{-4} . \]

According to the epidemiological data by Haukka et al. (2003), the estimated values for \( p \) and \( s \) are \( p = 1.29 \times 10^{-2} \) and \( s = 6.54 \times 10^{-3} \). Therefore, we have \( \nu = 1.76 \times 10^{-3} \) for the average mutation rate \((1.48 \times 10^{-4}) \), \( \nu = 1.76 \times 10^{-2} \) for the highest mutation rate \((1.48 \times 10^{-4}) \), and \( \nu = 1.76 \times 10^{-4} \) for a relatively low mutation rate \((1.48 \times 10^{-6}) \).

4.4 Implications for association studies of schizophrenia

We present here some applications of the P-criterion to association studies of schizophrenia. The results suggest that common disease/common variant hypothesis is unlikely to fit schizophrenia and that an enormous sample size is required to detect a nuclear susceptibility gene for schizophrenia.
4.4.1 Calculation of an upper bound of the effect size of a putative susceptibility gene of a given frequency

Using the P-criterion, we can calculate an upper bound of the effect size of a putative susceptibility gene of a given frequency.

Effects size of a susceptibility gene M is expressed by odds ratio defined as

\[
OR = \frac{m_A(1 - m_{ij})}{(1 - m_A)m_{ij}},
\]

which is monotonically increasing for \( m_A \) and monotonically decreasing for \( m_{ij} \). Since the criterion demands \( m_{ij} < m_A < m_{ij} + \nu \), we have

\[
\frac{m_{ij}(1 - m_{ij})}{(1 - m_A)m_{ij}} < \frac{(m_A + \nu)(1 - m_{ij})}{(1 - m_{ij} - \nu)m_{ij}}, \quad \text{or} \quad 1 < OR < 1 + \frac{\nu}{m_{ij}(1 - \nu - m_{ij})}
\]

for \( 0 < m_{ij} < 1 - \nu \).

And since the criterion demands \( m_A - \nu < m_{ij} \), we have

\[
1 < OR < \frac{m_A(1 - m_A + \nu)}{(1 - m_A)(m_A - \nu)} = 1 + \frac{\nu}{(1 - m_A)(m_A - \nu)}
\]

for \( 1 - \nu \leq m_{ij} < 1 \).

Thus, we have an upper bound of the effect size for a given frequency.

From the above we can easily see that the common disease/ common variant hypothesis, which proposes that common alleles at a handful of loci interact to cause a common disease, is unlikely to fit schizophrenia. No common alleles with population frequency between 0.05 and 0.95 can have large effects for schizophrenia: the odds ratio of every common risk allele is less than 1.04 for the average mutation rate, less than 1.58 for the highest mutation rate, and less than 1.004 for a relatively low mutation rate (Table 1).

4.4.2 Calculation of range of the frequency of a putative susceptibility gene of a given effect size

By solving the inequality (7) or (8), we can estimate the range of gene frequency for a given effect size. Thus, we can see that susceptibility genes of the average mutation rate and a moderate effect that meet the criterion are limited to ‘very rare variants’ or ‘very common variants’. For example, suppose \( \mu = 1.48 \times 10^{-5} \) and \( OR=5.0 \), then we have: \( \nu = 1.76 \times 10^{-2} \) and

\[
4 < \frac{\nu}{m_{ij}(1 - \nu - m_{ij})}.
\]

Solving this inequality, we get either \( 0 < m_{ij} < 0.00044 \) (that is, \( 0 < m_A < m_{ij} + 0.00176 < 0.0022 \)) or \( m_{ij} > 0.9977 \).
Impact of Epidemiology on Molecular Genetics of Schizophrenia

Table 1. Upper bounds of odds ratio for given allele frequencies in the unaffected population

<table>
<thead>
<tr>
<th>$m_u$</th>
<th>$\mu = 1.48 \times 10^{-6}$</th>
<th>$\mu = 1.48 \times 10^{-5}$</th>
<th>$\mu = 1.48 \times 10^{-4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>$&lt; 1.02$</td>
<td>$&lt; 1.18$</td>
<td>$&lt; 2.81$</td>
</tr>
<tr>
<td>0.02</td>
<td>$&lt; 1.009$</td>
<td>$&lt; 1.09$</td>
<td>$&lt; 1.92$</td>
</tr>
<tr>
<td>0.05</td>
<td>$&lt; 1.004$</td>
<td>$&lt; 1.04$</td>
<td>$&lt; 1.38$</td>
</tr>
<tr>
<td>0.1</td>
<td>$&lt; 1.002$</td>
<td>$&lt; 1.02$</td>
<td>$&lt; 1.20$</td>
</tr>
<tr>
<td>0.3</td>
<td>$&lt; 1.0009$</td>
<td>$&lt; 1.009$</td>
<td>$&lt; 1.09$</td>
</tr>
<tr>
<td>0.5</td>
<td>$&lt; 1.0008$</td>
<td>$&lt; 1.008$</td>
<td>$&lt; 1.08$</td>
</tr>
<tr>
<td>0.7</td>
<td>$&lt; 1.0009$</td>
<td>$&lt; 1.009$</td>
<td>$&lt; 1.09$</td>
</tr>
<tr>
<td>0.9</td>
<td>$&lt; 1.002$</td>
<td>$&lt; 1.02$</td>
<td>$&lt; 1.24$</td>
</tr>
<tr>
<td>0.95</td>
<td>$&lt; 1.004$</td>
<td>$&lt; 1.04$</td>
<td>$&lt; 1.58$</td>
</tr>
<tr>
<td>0.98</td>
<td>$&lt; 1.009$</td>
<td>$&lt; 1.10$</td>
<td>$&lt; 8.49$</td>
</tr>
</tbody>
</table>

4.4.3 Calculation of the required sample size and the power of an association study

Using the P-criterion we can calculate a lower bound of sample size required in an association study of a given power as well as an upper bound of the power of an association study of a given sample size.

Concerning the required sample size $2N$ ($N$ case-control pairs) and the power $1-\beta$ of an association study, we have the well-established formulae (Ohashi & Tokunaga, 2002):

$$N \geq \frac{1}{2} \left( \frac{z^*_a \sqrt{2x(1-x)} + z_\beta \gamma}{d} \right)^2,$$

and

$$1-\beta = \Phi \left( \frac{\sqrt{2Nd - z^*_a \sqrt{2x(1-x)}}}{\gamma} \right).$$

Here, $\Phi$, $z^*_a$, and $z_\beta$ denote the cumulative distribution function of the standard normal curve, the two sided $a$ point ($a$: a significant level) and the upper $\beta$ point of the standard normal curve, and $x$ (population frequency of the allele) and $\gamma^2$ are defined as follows.

$$x = \frac{1}{2} (m_A + m_U), \quad \gamma^2 = m_A(1 - m_A) + m_U(1 - m_U) = 2x(1-x) - \frac{1}{2}d^2$$

For the average mutation rate $\mu = 1.48 \times 10^{-5}$, we have $\nu = 1.76 \times 10^{-3}$. Suppose $0.0005 < x < 0.9995$, then we have $2x(1-x) > 0.9995 \times 10^{-3}$. From the P-criterion, we have:

$$\frac{1}{2}d^2 < \frac{1}{2} \nu^2 < 1.6 \times 10^{-6} < 2x(1-x) \times 0.002.$$

Therefore, we have the following approximation with an error smaller than 0.2%:
\[ \gamma^2 = 2x(1-x) - \frac{1}{2}d^2 \geq 2x(1-x) , \text{ or } \gamma \geq \sqrt{2x(1-x)}. \]

Thus, we have:

\[
N \geq \frac{1}{2} \left( \frac{z^*_{a} \sqrt{2x(1-x)} + z_{\beta/2}}{d} \right)^2 \geq \left( \frac{z^*_{a} + z_{\beta/2}}{d} \right)^2 x(1-x) \geq \left( \frac{z^*_{a} + z_{0.05}}{\sqrt{d \cdot 1.76 \times 10^{-3}}} \right)^2 x(1-x) = 3.72 \times 10^6
\]

for \( x = 0.5 \). Therefore, more than 3.7 million case-control pairs are required in a genome-wide association study with a power 0.95 to detect a susceptibility variant of the average mutation rate and a population frequency between 0.0005 and 0.9995.

Similarly we can see that more than 37,000 case-control pairs are required in a genome-wide association study with a power 0.95 to detect a susceptibility variant of the highest mutation rate (\( \mu = 1.48 \times 10^{-4} \)) and a population frequency between 0.005 and 0.995.

Finally, let us consider the case with a relatively low mutation rate \( \mu = 1.48 \times 10^{-6} \), which corresponds to \( \nu = 1.76 \times 10^{-4} \). In this case, more than 370 million case-control pairs are required in a genome-wide association study with a power 0.95 to detect a susceptibility variant of a population frequency between 0.000005 and 0.999995. Therefore it would take more than several hundred years to gather the required number of samples even if all of the affected individuals in the world were to be recruited to the study.

5. Mitochondrial DNA (mtDNA) hypothesis of schizophrenia

In this final section, we discuss on the nature of those schizophrenia-associated genes that do not meet the P-criterion, suggesting that these genes should be resistance genes that reduce the morbid risk and severity of the disease. We show that the results of association studies to date are compatible with the mitochondrial genome model but not with the nuclear genome model and propose a new hypothesis which assumes that the risk loci are in the mtDNA. We present eight major predictions of this hypothesis, and discuss that these predictions seem to accord with the other epidemiological findings and the results of the genetic and the pathophysiological studies to date.

5.1 Nature of schizophrenia-associated genes that do not meet the P-criterion

Now, let us consider the nature of those schizophrenia-associated genes that do not meet the persistence criterion. The inequality \( d \geq v \) implies \( s_{\mu} < s \), where \( s_{\mu} \) and \( s \) denote the selection
coefficient in the affected subpopulation with an allele M and in the affected population respectively. Therefore, such genes, if sustained by mutation-selection balance, cannot be susceptibility genes but resistance genes that reduce severity and risk of the disease (see 4.2). If they were not resistance genes, their frequencies in the affected population must have been reduced to the same level in the unaffected population.

5.2 The results of association studies to date accord with the mitochondrial genome model but not with the nuclear genome model

Since a resistance gene in the nuclear genome model cannot be associated with the disease unless it is linked with a susceptibility gene, resistance genes in the nuclear model should be located in the vicinity of susceptibility genes, which disagrees with the results of association studies to date.

For example, on the chromosome 1, all of the schizophrenia-associated genes that could meet the criterion (RGS4, PLXNA2, DISC1) are located on 1q, while four resistance genes (MHTFR, GRIK3, PDE4B, GSTM1) are on 1p (Table 2). Fifteen resistance genes are located on 2q, 5q, 7q, 10q, 11p, 12p, 12q, 13p, 13q, 16p, 17p, and 19q, where no schizophrenia-associated variants that could meet the criterion are located (data: not shown). Therefore, the results of association studies to date argue against the nuclear genome model.

A possible interpretation which accords with the nuclear genome model might be that many nuclear susceptibility genes of less than the highest mutation rates have not been detected by association studies to date due to lack of power. In this case, however, an enormous sample size (more than 3.7~370 million case-control pairs) would be required to identify them as was mentioned above. In other words, such an enormous sample size is required to prove the nuclear genome model.

On the other hand, every resistance gene on any chromosome can be associated with schizophrenia in the mitochondrial genome model; since mtDNA is transmitted only via females and there is no link between the nuclear genome and the mitochondrial genome, every nuclear genome which interacts with a pathogenic mitochondrial genome to alter severity and risk of the disease is subject to natural selections in the predisposed maternal lineage that succeeds to a same pathogenic mitochondrial genome. Therefore, every resistance gene for schizophrenia in the mitochondrial genome model is to be subject to a positive selection in the predisposed maternal lineage, thereby associating with schizophrenia.

Thus, the mitochondrial genome model is compatible with the results of the association studies to date.

It should be noted that in the mitochondrial genome model every facilitating gene (a gene that increases the severity and morbid risk in the predisposed population) on any chromosome may diminish in the predisposed matrilineal pedigrees by negative selection, thereby negatively associating with the disease.

Schizophrenia-associated variants listed in the top 45 in the SZGene Database (the version of 10th December, 2010) were selected. Based on the genotype distributions in meta-analyses, allele frequencies and the case-control differences were calculated. 4 variants at the 3 loci (RGS4, PLXNA2, DISC1) could meet the criterion under the assumption that the mutation
rates at those loci are near the upper limit in the autosomes. All of them are located on 1q, while 4 resistance genes (MHTFR, GRIK3, PDE4B, GSTMI) are on 1p. * schizophrenia-associated alleles; variants that could meet the criterion are shown in bold characters

<table>
<thead>
<tr>
<th>Genes and SNPs</th>
<th>Location</th>
<th>Allele (minor/major)</th>
<th>mA</th>
<th>mU</th>
<th>OR</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>MHTFR</td>
<td>1p36.22</td>
<td>T*/C</td>
<td>0.3532</td>
<td>0.3211</td>
<td>1.15</td>
<td>0.032</td>
</tr>
<tr>
<td>GRIK3</td>
<td>1p34.3</td>
<td>G*/T</td>
<td>0.2600</td>
<td>0.2226</td>
<td>1.25</td>
<td>0.037</td>
</tr>
<tr>
<td>rs910694</td>
<td>1p31.3</td>
<td>C/T*</td>
<td>0.5780</td>
<td>0.5477</td>
<td>1.30</td>
<td>0.030</td>
</tr>
<tr>
<td>GSTM1</td>
<td>1p13.3</td>
<td>ins-allele/del-allele*</td>
<td>0.7546</td>
<td>0.7140</td>
<td>1.35</td>
<td>0.041</td>
</tr>
<tr>
<td>RGS4</td>
<td>1q23.3</td>
<td>A/G</td>
<td>0.4920</td>
<td>0.4744</td>
<td>1.08</td>
<td>0.0176</td>
</tr>
<tr>
<td>rs2661319</td>
<td>1q32.1</td>
<td>G*/A</td>
<td>0.3056</td>
<td>0.2657</td>
<td>1.42</td>
<td>0.040</td>
</tr>
<tr>
<td>PLXNA2</td>
<td>1q32.2</td>
<td>A/G</td>
<td>0.8434</td>
<td>0.8001</td>
<td>1.32</td>
<td>0.043</td>
</tr>
<tr>
<td>G*/C*</td>
<td>0.92840</td>
<td>0.91243</td>
<td>0.32</td>
<td>0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISCI</td>
<td>1q42.3</td>
<td>A*/G</td>
<td>0.03069</td>
<td>0.01735</td>
<td>1.80</td>
<td>0.013</td>
</tr>
<tr>
<td>rs999710</td>
<td>A*/G</td>
<td>0.3989</td>
<td>0.3819</td>
<td>1.07</td>
<td>0.0170</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Schizophrenia-associated genes on the chromosome 1 that could meet the criterion

5.3 The mtDNA hypothesis for schizophrenia and its predictions

Thus we propose here a new hypothesis which insists that the risk loci for schizophrenia are in the mtDNA.

Mitochondria are involved in a variety of major cellular events such as oxidative phosphorylation, free radical production and Ca\(^{2+}\) buffering, and play an active role in apoptosis. They possess two classes of antioxidant defense system (non-enzymatic and enzymatic), and structurally and functionally intact mitochondria serve as a net sink rather than a net source of reactive oxygen species (ROS) (Andreyev et al., 2005). ROS-defenses are severely undermined in structurally compromised mitochondria (Andreyev et al., 2005). Thus, mitochondrial dysfunction, presumably through imbalance of ROS production and removal (Andreyev et al., 2005), raises ROS emission (Esposito et al., 1999; Senoo-Matsuda et al., 2001) and causes intracellular oxidative stress.

Because abnormal mtDNA may cause mitochondrial dysfunction, the hypothesis predicts: (1) enhanced oxidative stress and disturbed energy metabolism in predisposed individuals, which may cause various pathogenic alterations such as genomic instability, aberrations in neurodevelopment, and the brain dysfunction. Furthermore, because mtDNA can be transmitted only through females and there is no link between the nuclear genome and the
mitochondrial genome, the mtDNA hypothesis predicts: (2) a higher maternal transmission of schizophrenia, and (3) positive associations between resistance genes and schizophrenia as well as negative associations between facilitating genes and schizophrenia (see 5.2). These predictions seem to be consistent with other major epidemiological findings and the results of the genetic and the pathophysiological studies to date.

5.3.1 Mitochondrial dysfunction and enhanced oxidative stress in affected individuals

The hypothesis predicts that patients with schizophrenia show mitochondrial dysfunction and enhanced oxidative stress.

Indeed, in the past decade, mitochondrial dysfunction and oxidative stress in schizophrenia has been suggested by several independent lines of evidence (for review, see Marchbanks et al., 1995; Ben-Shaffer, 2002; Wood et al., 2009); those include mitochondrial hypoplasia, disturbed oxidative phosphorylation, and altered mitochondrial-related gene expression in several cell lines.

The pioneering works in this field may be noteworthy (Utena & Niwa, 1992). As early as 1950, Hayashi, in a longitudinal study on glucose metabolites in blood sampled from the superior bulb of the internal jugular vein of schizophrenics, observed a decreased carbonic dioxide production in the brain and a higher level of lactate and glutathione, the brain’s dominant free radical scavenger, in patients in an acute exacerbation of the illness. Utena and Ezoe (1951) reported a decreased glucose consumption in vitro in cortical brain tissues sampled from patients with schizophrenia who underwent prefrontal leukotomy. Takahashi (1953) confirmed this finding and emphasized the necessity of further investigations on oxidative phosphorylation in the brain tissue of schizophrenics. In line with those findings was the report by Stabenau et al. (1969), who observed, in a biochemical study of discordant monozygotic twin pairs, that lactate production and the lactate-pyruvate ratio were higher in the affected twins than the unaffected cotwins. More recently, Prabakaran et al. (2004), in a large-scale functional genomics study, suggested a state of intermittent or chronic hypoxic stress and mitochondrial dysfunction in the brain of patients with schizophrenia.

5.3.2 The mode of transmission

The hypothesis predicts a higher maternal transmission of schizophrenia. Although there has been no convincing evidence for maternal transmission of schizophrenia, several reports suggest a higher maternal transmission of schizophrenia (Shimizu et al., 1987; Goldstein et al., 1990; Valero et al., 1998; Li et al., 2007).

Some researchers have proposed the hypothesis that schizophrenia is associated with \textit{de novo} mutations arising in paternal germ cells (Malaspina et al., 2001; Zammit et al., 2003; Byrne et al., 2003; El-Saadi et al., 2004; Sipos et al., 2004). It is based on the observation (‘paternal age effect’) that the risk of schizophrenia in the offspring seems to increase as the paternal age advances from 20 years to over 50 years.

However, the difference in the mean ages of fathers between affected and unaffected individuals are not very large ($< 1.7$ years) (Malaspina et al., 2001; El-Saadi et al., 2004). Furthermore, the risk of schizophrenia was also increased in the offspring of younger men.
(< 21 years) (Malaspina et al., 2001; El-Saadi et al., 2004; Sipos et al., 2004) as well as in the offspring of younger women (< 20 years) (El-Saadi et al., 2004). Therefore, major roles of paternally derived mutations in schizophrenia seem to remain unsubstantiated.

Indeed, no available data can exclude the possibility that the ‘paternal age effect’ has a ‘maternal origin’; while women in many countries today may be usually supposed to bear children after the age of 20 years and not to marry much older men or too young men unless the men have special socio-economic benefits, a certain proportion of predisposed women might behave differently.

It should be noted that in the famous twin study by Gottesman and Bertelsen (1989) which included almost equal number of male and female monozygotic twins, most schizophrenic twins whose offspring are affected are females (12 out of 14), implying that the transmission was mainly via females ($p = 14C_{12} \times 0.5^{14} + 14C_{13} \times 0.5^{14} + 13C_{14} \times 0.5^{14} < 0.007$). While this gender effect might be due to non-genetic factors such as stronger psychological interactions between mother and child, we must also consider the possibility that it is due to the closer genetic relationship between mother and child, i.e. the mtDNA.

5.3.3 Sex difference and a protective effect of estrogen in schizophrenia

The hypothesis predicts that endogenous antioxidants exhibit a protective effect against schizophrenia, and may give a plausible explanation for sex difference of the disease.

A consistent and specific finding for schizophrenia is that the age at onset is significantly lower in males than in females (Jablensky, 1995; Kulkarni & Fink, 2000); schizophrenia starts earlier on average in males and reaches its peak between 15 and 25 years of age, whereas in females it occurs almost between 20 and 30 years of age and shows a less steep curve after that age. It also appears that women are vulnerable to relapses or first episode of schizophrenia in the perimenopausal period (the second peak of onset for females) (Kulkarni & Fink, 2000), when estrogen production diminishes. A close association between premenstrual or menstruation phase and exacerbation of the illness in females has been well documented (Kulkarni & Fink, 2000). In addition, less negative symptoms, less brain morphological changes, and better response to neuroleptic medication are relatively consistent finding in female patients with schizophrenia (Jablensky, 1995; Goldstein & Lewine, 2000).

These observations lead to the concept that estrogen protects predisposed females (Kulkarni & Fink, 2000), which seems to accord with the hypothesis; estrogen has been shown to have antioxidant activity due to its intrinsic antioxidant structure that lies in the phenolic moiety of the steroidal compound (Behl, 2002), to increase antioxidant enzyme activities (Strehlow et al., 2003; Pajović et al., 2003), and to have neuroprotective effect against oxidative stress (Behl, 2002; Brann et al., 2007). Furthermore, mitochondrion has estrogen binding sites (Monje & Boland, 2001; Chen et al., 2004) and estrogen increases mitochondrial efficiency and reduces intracellular oxidative stress (Stirone et al, 2005).

5.3.4 Low comorbidity between schizophrenia and rheumatoid arthritis

The hypothesis predicts that diseases predisposed by facilitating genes, if present, would be negatively associated with schizophrenia. Rheumatoid arthritis could be one of such candidates.
A role of oxidative stress in the pathogenesis of rheumatoid arthritis has been suggested by several lines of evidence (for review, see Hitchon et al., 2004). In addition, it has been shown that chronic oxidative stress in the synovial T lymphocytes is not secondary to exposure to environmental free radicals but originates from intracellularly produced reactive oxygen species (Remans et al, 2005). Therefore, a presumptive susceptibility gene for rheumatoid arthritis, which may cause intracellular oxidative stress in several cell lines, could be a facilitating gene for schizophrenia in this model and is likely to be subject to a negative selection in the predisposed matrilineal pedigrees.

Indeed, robust evidence shows a negative association between schizophrenia and rheumatoid arthritis while the exact mechanism is still unknown (Vinogradov et al., 1991; Jablensky, 1995; Rubinstein, 1997; Oken 1999). According to the nuclear genome model, several hypotheses have been proposed such that pathogenic genes for schizophrenia may be protective genes for rheumatoid arthritis and vice versa.

Thus the mitochondrial genome model may offer a new explanation for the low comorbidity between schizophrenia and rheumatoid arthritis and the additional prediction: most of patients with both of the diseases would be females because the survival rate of males in early life stage must be remarkably reduced due to lack of the antioxidant defense by estrogen, and show more negative symptoms, poorer response to neuroleptic medication, and/or more morphological changes in the brain.

5.3.5 Prenatal risk factors for schizophrenia

The hypothesis predicts that early-life exposure to environments which induce strong oxidative stress can increase the risk of later development of schizophrenia in the predisposed population.

Indeed, prenatal environmental factors such as severe nutritional deficiency (Susser, et al., 1996), exposure to increased homocysteine (Brown et al., 2007) or lead (Opler & Susser, 2005), and infection of influenza virus (Limosin et al., 2003; Brown et al., 2004; Opler & Susser, 2005) and Toxoplasma gondii (Brown et al., 2005) have been suggested to increase the risk for schizophrenia. More recently, it has been suggested that central nervous system infections of cytomegalovirus or mumps virus in childhood may also increase the risk for schizophrenia (Dalman et al., 2008). All of these factors have been shown to affect mitochondria, inducing strong intracellular oxidative stress and/or apoptosis (Akaike et al., 1990; Edlund et al., 1994; Speir et al., 1998; He et al., 2003; Berger et al., 2004; Zaki et al., 2005; Gupta et al., 2004; Krumen et al., 2006; Wang et al., 2006; Poncet et al., 2006; Chang et al., 2007; Nishikawa et al., 2007).

5.3.6 Increased obstetric complications in the birth of patients with schizophrenia

It has been suggested that mitochondrial dysfunction may be involved in the etiology of preeclampsia (Shanklin et al., 1990; Barton et al., 1991; Furui et al., 1994). In addition, a high incidence of preeclampsia, eclampsia, and stillborn infants has been observed in a family with a known mitochondrial disorder (Torbergsen et al. 1989). Folgero et al. (1996) demonstrated two separate mtDNA point mutations in two families having a high incidence of preeclampsia and eclampsia.
Therefore, the hypothesis predicts that the risk of preeclampsia, eclampsia, or stillbirth may be increased in the birth of patients with schizophrenia as well as in the pregnancies of women with schizophrenia. Indeed, an excess of stillbirths and neonatal deaths among women with schizophrenia has been reported by several investigators (Sobel, 1961; Rieder et al., 1975; Modrzewskia, 1980; Webb et al., 2005).

Furthermore, there has been a body of evidence for an increased risk of obstetric complications in the birth of patients with schizophrenia (Dalman et al., 1999; Cannon et al., 2002). A meta-analysis of population-based data (Cannon et al., 2002) found significant estimates for three main categories of obstetric complications: (1) complications of pregnancies, (2) abnormal fetal growth and development, and (3) complications in delivery. Among all, preeclampsia was the strongest individual risk factor detected in the largest single population-based cohort study to date (Dalman et al., 1999).

Although obstetrical events in schizophrenia are often considered as having a direct causative effect, none of the available data can refute the hypothesis that they are merely markers of some other causal process (Rapaport et al., 2005), such as mitochondrial dysfunction which is implicated in this hypothesis.

5.3.7 An apparent signature of positive selection in schizophrenia-associated genes

Since the positive selection of the schizophrenia-associated alleles mentioned above occurs only in the predisposed matrilineal pedigrees, a ubiquitous subpopulation in humans, frequencies of those alleles may not be so high in the general population as if the selection had occurred recently in the general population.

Thus, the hypothesis predicts that every schizophrenia-associated nuclear gene shows an apparent signature as if it had been subject to a positive selection in the recent evolutionary history of humans. Recent two reports (Lo et al, 2007; Crespi et al., 2007) seem to be in line with this prediction.

On the other hand, the nuclear genome model predicts that every schizophrenia-associated nuclear gene shows an apparent signature of negative selection due to the strong negative selection pressure.

5.3.8 Genomic instability

It is generally thought that a major cause of DNA damage that leads to mutations is reactive oxygen species, which are generated as a normal part of oxygen metabolism but are also produced by ionising radiation, metabolism of exogenous compounds (Hussain et al., 2003; Finkel, 2003). It has been shown that endogenous mitochondrial oxidative stress can induce many types of DNA damage including double strand breaks, end-to-end fusions, base and sugar modifications, DNA-protein cross-links, and gross chromosomal rearrangements (Ragu et al., 2007; Samper et al., 2003).

Therefore, the hypothesis predicts that the enhanced oxidative stress may cause genomic instability during meiosis and/or early phase of ontogeny, producing increased rates of random point mutations and/or structural variants of the nuclear genome in the
predisposed population. In addition, genomic instability may be more pronounced in males due to lack of antioxidant protection by estrogen.

There have been numerous reports of associations between schizophrenia and chromosomal abnormalities including fragile sites, reciprocal translocations, inversions, insertions, deletions, disomy and trisomy in many autosomes, and sex chromosome aneuploidies (Macintyre et al, 2003). However, with an exception of 22q11 deletion, none of these have been consistently replicated, and with another exception of (1,11) (q42;q14.3) balanced translocation, none provides convincing evidence for the location of a ‘susceptibility’ gene (Kirov et al., 2005).

A popular explanation in the nuclear genome model may be that most of these structural variants are coincidental findings of no clinical significance. Alternatively, those alterations may indicate genomic instability in schizophrenia. An increased risk of schizophrenia in individuals with 22q11 deletion (Pulver et al., 1994; Murphy et al., 1999) might be due to haplodeficiency of presumptive resistance genes of gain-of-function type and/or presumptive facilitating genes of loss-of-function type aggregated on 22q11.

More recently, it has been reported that rare structural variants such as microdeletions or microduplications of sizes ranging from 100kb to 15MB throughout the genome are more frequent among individuals with schizophrenia than unaffected individuals (Walsh et al., 2008). While many of those structural variants duplicate or delete genes in neurodevelopmental pathways, one third of those do not disrupt genes, leaving their role in causation of the disease unwarranted. Another recent report (Xu et al., 2008) has shown that de novo copy number mutations are increased in sporadic schizophrenia. However, the cytobands of those copy number mutations are diverse among the affected individuals and their roles in the pathogenesis still remain unclear. Therefore, no available data can refute the possibility that those structural variants and copy number mutations are not the causes of schizophrenia but the results of the genomic instability in schizophrenia predicted by our hypothesis.

Indeed, direct measure of the de novo mutation rates shows an increased mutation rate in schizophrenia (Awadalla et al., 2010), and genomic and epigenomic instability has been suggested in schizophrenia (Smith et al., 2010). Furthermore, it has been shown that blood cells from patients with schizophrenia present a higher rate of folate-sensitive fragile sites, and that male patients exhibit twice as many fragile sites as females while there are no age effects (Demirhan et al., 2006). This sex difference may indicate that increased fragile sites expression (genomic instability) is the results of enhanced oxidative stress in patients with schizophrenia.

6. Conclusion

Genetic research of schizophrenia based on the nuclear genome model has been one of the most active areas in psychiatry for the past two decades. Although this effort is ongoing, results of association studies have been inconsistent and the situation of molecular genetics of schizophrenia today has become much confused just contrary to our expectation. The consistent major epidemiological findings of schizophrenia, coupled with the results of association studies to date, argue against the nuclear genome model for schizophrenia.
Rather, they seem to argue in favor of the mitochondrial genome model, suggesting a necessity of paradigm shift from the nuclear genome model to the mitochondrial genome model in genetic research of schizophrenia in the coming years.

**Note:** Cross-generational reduction of females with pathogenic genes in the mitochondrial genome model

At first we define several notations. \( N_1 \): the number of normal females in the first generation; \( N_2 \): number of female offspring of normal females; \( S_1 \): the number of unaffected female siblings of patients in the first generation; \( S_2 \): the number of female offspring of unaffected female siblings of patients; \( P_1 \): the number of female patients; \( P_2 \): the number of female offspring of female patients; \( r \) (0 < \( r \) < 1): the proportion of gene carriers in normal females in the first generation. Then the number of female gene carriers in the first generation is \( rN_1 + S_1 + P_1 \) and the frequency of female gene carriers in the first generation is given by:

\[
f_1 = \frac{rN_1 + S_1 + P_1}{N_1 + S_1 + P_1} = r + \frac{S_1 + P_1}{N_1 + S_1 + P_1} \cdot (1 - r).
\]

And the frequency of female gene carriers in the second generation is given by:

\[
f_2 = \frac{rN_2 + S_2 + P_2}{N_2 + S_2 + P_2} = r + \frac{S_2 + P_2}{N_2 + S_2 + P_2} \cdot (1 - r).
\]

Thus we have (Table 3):

\[
-\Delta = f_1 - f_2 = \left( \frac{S_1 + P_1}{N_1 + S_1 + P_1} - \frac{S_2 + P_2}{N_2 + S_2 + P_2} \right) \cdot (1 - r) < 5.06 \times 10^{-3}.
\]

<table>
<thead>
<tr>
<th>( N )</th>
<th>( S )</th>
<th>( P )</th>
<th>Total</th>
<th>( (S+P) / \text{Total} )</th>
</tr>
</thead>
<tbody>
<tr>
<td># of females</td>
<td>410,093</td>
<td>11,873</td>
<td>4,784</td>
<td>426,750</td>
</tr>
<tr>
<td># of female children</td>
<td>366,460</td>
<td>10,969</td>
<td>1,917</td>
<td>379,346</td>
</tr>
<tr>
<td>(-\Delta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Epidemiological data by Haukka et al. (2003)

In this largest-sampled cohort study to date, Haukka et al. comprised all births in Finland during 1950-1959 (\( N = 870,093 \)) and followed up through the National Hospital Discharge Register for Hospitalizations between 1969 and 1992. \( N \): normal females; \( S \): unaffected female siblings of patients; \( P \): female patients with schizophrenia

**References**


formation and apoptosis in rod photoreceptors of lead-exposed mice. Proc Natl Acad Sci USA, Vol. 100, No. 3, (Feb 2003), pp. 1022-1027, ISSN 0027-8424


www.intechopen.com


Impact of Epidemiology on Molecular Genetics of Schizophrenia


This book represents an overview on the diverse threads of epidemiological research, brings together the expertise and enthusiasm of an international panel of leading researchers to provide a state-of-the-art overview of the field. Topics include the epidemiology of dermatomycoses and Candida spp. infections, the epidemiology molecular of methicillin-resistant Staphylococcus aureus (MRSA) isolated from humans and animals, the epidemiology of varied manifestations neuro-psychiatric, virology and epidemiology, epidemiology of wildlife tuberculosis, epidemiologic approaches to the study of microbial quality of milk and milk products, Cox proportional hazards model, epidemiology of lymphoid malignancy, epidemiology of primary immunodeficiency diseases and genetic epidemiology family-based. Written by experts from around the globe, this book is reading for clinicians, researchers and students, who intend to address these issues.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
