NeuRA Discover. Conquer. Cure.

SCHIZOPHRENIA LIBRARY

Genetics overview

The Building Blocks

Chromosomes

Chromosomes are strings of deoxy-ribonucleic acid (DNA) which form the human genetic material. In general, humans have 46 chromosomes, arranged into 22 pairs of chromosomes (known as autosomes) plus two sex chromosomes, X or Y. Females have two X chromosomes and males have an X and Y chromosome. Each pair of chromosomes are known as 'homologous' chromosomes. One set of 23 chromosomes comes from the mother's egg, and the other set are from the father's sperm.

DNA

The DNA molecule is made up of subunits called nucleotides. Each nucleotide contains one of four base molecules: adenosine (A), thymine (T), guanine (G), or cytosine (C). The base, together with a sugar and a phosphate, forms a nucleotide subunit. Many nucleotides are joined together into long strands. The DNA molecule is two strands of nucleotides in a "double-helix" arrangement, which has a ladder appearance. The bases are paired to form the rungs of the ladder, so that A always pairs with T, and C always pairs with G. These double-helix strands of DNA form the chromosomes.

Genes

Each chromosome is made up of sections, known as genes. A gene is a specific sequence of A, T, C and G bases, and each gene is 'coded' by a unique sequence. The combination of bases that make up each gene will influence the function of the gene. Genes produce different proteins, depending on the sequence of bases. Each protein is specific to that gene, and will be involved in different aspects of growth, development, and cellular function.

A person's unique traits and features are passed across generations through genes. One copy of each gene is located on each chromosome, so that each person has two copies of every gene; one from each parent. An exception to this rule is on the sex chromosomes, where males have one X and one Y, which are not identical, but females have two X chromosomes.

Polymorphisms

Within a population of people, there may be several different forms of the same gene. These variant forms are called alleles. One person may carry two different alleles of the same gene, one from each parent. Some diseases find that one allele of the gene has greater risk for the disease, so that people with that allele ('carriers') are at greater risk of the disease. The variations in DNA sequence between the different alleles of any gene are known as polymorphisms. A variation in base sequence at a particular point in the DNA is a polymorphism. A gene might have several polymorphisms across a population of people, resulting in several alleles. Many alleles differ in DNA sequence by only a single nucleotide polymorphism, or a SNP. A polymorphism in a gene might produce a slightly different protein, or it could even cause the gene to have a different function altogether. As a result, polymorphisms are one important source of variation between people within a population. The combination of gene polymorphisms in each person's DNA will be unique to

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

that individual. Polymorphisms are identified by a unique 'rs' number, which stands for Reference SNP. This is the universal nomenclature for identifying polymorphisms.

Genetics in schizophrenia

Often, when a gene is associated with an illness, "typical" inheritance patterns mean that a person carrying that gene (or the affected allele) will be affected by the illness, and this gene may be passed down to any offspring. However, schizophrenia is appearing to be much more complex. Hundreds of genes have been linked to schizophrenia and these genes do not appear to follow typical patterns of inheritance across generations. Moreover, the additional influence of environmental factors significantly complicates the identification of susceptibility genes. This suggests that a susceptibility to schizophrenia may involve associations with multiple genes which interact with each other, and with the environment, to result in the symptoms of schizophrenia.

Schizophrenia Gene Database

The sheer volume of information available has lead to the development of a website dedicated to compiling all the existing evidence for schizophrenia susceptibility genes. SchizophreniaGene (www.szgene.org) aims to collate the evidence for all genes and polymorphisms that have been implicated in schizophrenia, performing meta-analyses where possible to assess the strength of the evidence for each gene and its polymorphisms[1].

Homepage

From the entry page (Homepage), the database can be searched by chromosome, gene, protein, polymorphism, or by individual study. Searching by chromosome will provide links to all known susceptibility genes on that chromosome. Searching by gene takes the user directly to an information page specific to that particular gene. Searching by study takes the user to the information page for the gene contained in that study. Similarly, searching by protein takes the user to the information for the gene coding that protein. Searching by polymorphism takes the user directly to the meta-analysis performed on the evidence for that polymorphism, where available.

Search by chromosome

Selecting any of the chromosomes listed on the homepage will display a schematic diagram of the chromosome in question. This demonstrates all the genes on each chromosome that have been associated with schizophrenia. Clicking on any of the genes takes the user to each gene's information page. The MT chromosome is referring to the circular mitochondrial DNA, found within the mitochondria of a cell (distinct from the nucleus).

For an example, see the search results for Chromosome 6

Search by gene

Searching by gene takes the user to the relevant gene's information page, which is headed by the gene name, the chromosome on which this gene is located, and the protein it codes. A table summarises all of the studies investigating this gene, including details for the ethnicity of the sample population, the number of

subjects assessed and their ages, and whether the study reported a positive association of the gene in question with schizophrenia or a negative result. A positive association suggests the polymorphism is reported with increased or decreased frequency in people with schizophrenia. A negative result suggests there is no evidence for any significant association with schizophrenia. Some studies report a 'trend' association, which refers to a positive association that has not quite reached an acceptable level of statistical significance.

For an example, see the search results for the PGBD1 gene

Meta-analysis

Meta-analysis is a statistical technique that is used to pool the results of many studies on the same topic into one analysis, giving greater insight into the overall picture. Meta-analysis is important in genetic analyses, as an individual study may be affected by many external factors such as the ethnicity of a population, but combining many studies into a meta-analysis allows an assessment of the overall association of a gene with schizophrenia. The SchizophreniaGene website contains a link immediately above the gene overview table which takes the user to the web page containing a meta-analysis for each gene. Meta-analyses have been performed by the authors of the website where at least four studies are available which compare schizophrenia cases with healthy control subjects from the general population.

For an example, see the meta-analysis page of the PGBD1 gene

The meta-analysis tables are labelled according to the polymorphism's unique rs number. The first table on this page lists the allele frequencies for the gene in question. The allele frequency simply represents the proportion of copies of this gene which are accounted for by each individual allele, in other words, what percentage of people express each allele variant, as a proportion of the total. Reading the table across, the frequencies in each category "alleles" and "genotypes" will sum to 1.0 (100%). The 'allele' category considers each allele separately, while the 'genotype' category considers the combination of alleles (because each person carries two copies of every gene, one on each homologous or 'paired' chromosome).

Meta-analysis is used to combine the available evidence for each polymorphism, calculating overall how strongly it is linked to schizophrenia. This value is represented in the second figure on the meta-analysis page – the Forest plot. This figure shows the overall combined values (effect size) in **bold** type at the top of the list, and the data for each individual study below. The measure of effect size most commonly used to assess the association between the polymorphism and schizophrenia is called the odds ratio (OR). This value represents the 'odds' (probability) that the polymorphism in question is expressed more, or less, frequently in people with schizophrenia compared to the general population. An OR greater than one (> 1) suggests that people with schizophrenia have increased odds of carrying the polymorphism in question compared to healthy controls, in other words, the allele's frequency is increased in schizophrenia. An OR less than one (< 1) suggests that people with schizophrenia are less likely to be carrying the polymorphism in question compared to healthy controls. The reported confidence intervals indicate that 95% of all sample populations tested should fall within this interval, and is a measure of the reliability and precision of the effect size. I^2 is a measure of the consistency of the evidence – the variance (heterogeneity) within the sample being assessed. The I^2 value is the percentage of the variability that is true variation, as opposed to random chance.

NeuRA
Discover. Conquer. Cure.

Genetics overview

SCHIZOPHRENIA LIBRARY

Top genes

This website has identified the Top Genes which have the strongest evidence supporting their involvement in schizophrenia. This classification is based on the Human Genome Epidemiology Network (HuGENet) guidelines for assessing genetic association studies. There are three key criteria for assessment, including the amount of evidence available, the consistency of this evidence, and the amount of potential bias affecting the evidence.

The three criteria for assessment are used to determine the overall **quality** of the evidence. The first criterion is the amount of evidence available, and essentially refers to the sample size, and the number of alleles incorporated in this group of people. The second guideline is the consistency of replication, and refers to the amount of variance between the studies which contribute to the overall effect. It is assessed as I^2, which is the percentage of the variability that is true, rather than sampling error (chance). I^2 is graded using benchmark values: $I^2 < 25\%$ is grade A; I^2 between 25-50% is grade B; and $I^2 > 50\%$ is grade C. The third criterion considers the effect of plausible biases, including bias resulting from a low effect size; publication bias; or from a loss of significance after certain studies are excluded.

Overall evidence quality is graded as "A" (strong) if the assessment criteria receive three A grades, "B" (moderate) if they receive at least one B grade but no C grades, and "C" (weak) if they receive a C grade in any of the three assessment criteria. For more details see Top Results Methods.

For the most up to date information on the top genes associated with schizophrenia, please refer to the SzGene website (hyperlink above). The list of top genes includes those genes which show a significant summary effect in the analysis of all studies, or in analyses limited to samples of a specific ethnicity (e.g. Caucasian). Genes are ranked based on the gene variant with the best overall grade. For genes with identical grades, the ranking is based on *p*-value; and for genes with identical grade & *p*-value, ranking is based on effect size (OR).

Gene function

For useful information regarding the normal function of all identified genes, there are two National Centre for Biotechnology Information (NCBI) initiatives available, including the Entrez gene database (http://www.ncbi.nlm.nih.gov/gene/) and OMIM database (Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/sites/entrez?db=omim), which provide useful information regarding the normal function of all identified genes, so that each gene can be searched individually for an explanation of its function and expression patterns in humans. Note, these explanations are not presented in the context of schizophrenia and only report on the normal function of the gene. Understanding the full consequences of altered gene function is an area of continuing research. Ensure when viewing the search results that only genes for Homo sapiens are viewed, as other species are also contained within the Entrez database.

References

1. Allen NC, Bagade S, McQueen MB, Ioannidis JP, Kavvoura FK, Khoury MJ, Tanzi RE, Bertram L. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. *Nature Genetics*. 2008; **40**(7): 827-34.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au