Decoding the genetics of synaesthesia using state-of-the-art genomics

Sarah A. Graham[1], Katerina S. Kucera[1], Duncan A. Carmichael[3,4], Julian E. Asher[5], Allen Chan[5], Laura Murphy[5], Simon Baron-Cohen[5], Julia Simner[3], Simon E. Fisher[1,2]

[1] Language and Genetics Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands.
[2] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands.
[3] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK.
[4] Institute for Adaptive and Neural Computation, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK.
[5] Autism Research Centre, University of Cambridge, 18b Trumpington Road, Cambridge CB2 8AH, UK.

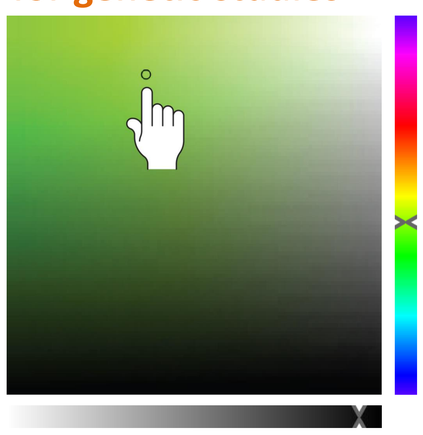
MAX PLANCK INSTITUTE FOR PSYCHOLINGUISTICS

Synaesthesia as a genetic trait

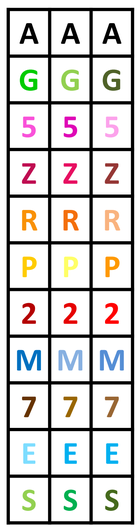
It was first noted over a century ago that synaesthesia runs in families. This suggests that genetic factors play an important role in an individual's likelihood of developing synaesthesia. However, specific genetic variants that predispose to synaesthesia have yet to be discovered, and the genetic foundations are likely to be complex and to vary among individuals. We are using state-of-the-art genome-wide approaches to investigate the genetic basis of synaesthesia.

Characterizing synaesthesia for genetic studies

For genetic studies involving large numbers of individuals, it is important to be able to assess the trait of interest in a quick and consistent way. For this reason, we focus on *grapheme-colour synaesthesia*, and we ask everyone who takes part in our research to complete the Synaesthesia Battery. This includes tasks such as the Colour Picker Test (right), which gives a score based on consistency of grapheme-colour associations.



The Colour Picker Test
http://synesthete.org



Human genetic variation

DNA is a long chain composed of four different kinds of *nucleotide*: A, T, C and G. The human *genome* contains 3 billion nucleotides. Even though 99.9% of the genome is the same across the whole human population, every individual has millions of differences (*variants*) compared to the *reference sequence*.

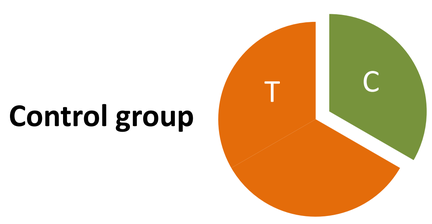
Reference sequence ...ATTCGCCTCGTTTTAAGACAAGTG...
Variant ...ATTCGCCTCGTTCTAAGACAAGTG...

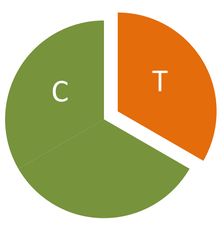
Many of these nucleotide differences are *common variants*, often called *single-nucleotide polymorphisms* or *SNPs*, which means that at least 1 person in 100 has a different nucleotide to the reference sequence. There are also *rare variants*, which are found in less than 1% of the population, and possibly only in one individual or their family.

We are using different technologies to look at both *common variants* and *rare variants*, to assess their role in synaesthesia.

The role of common genetic variants: association studies

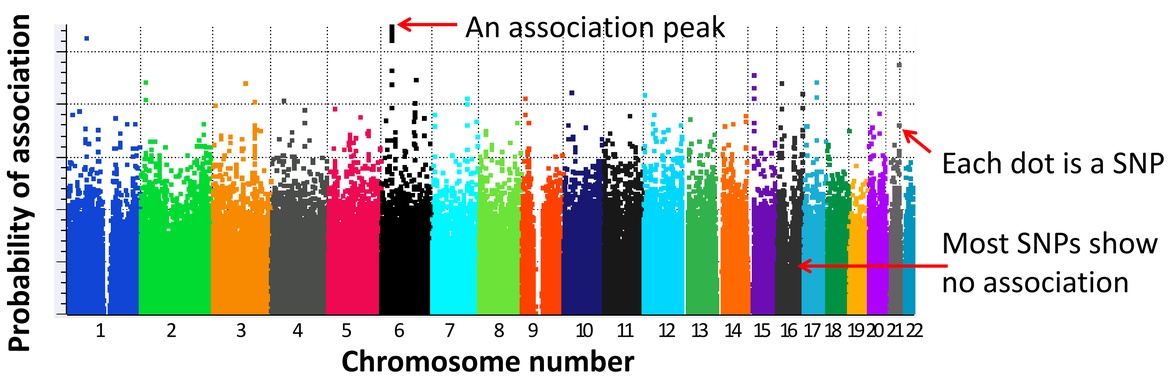
By statistically comparing the frequencies of the two different versions of a *common SNP* in a group of people who have a particular trait (eg. synaesthesia) with a group of people who don't (the control group), we can determine if the region of the genome around the SNP may be involved in the trait. This is called an *association* test. In the illustration below, the C variant is present significantly more frequently in the test group than the control group, so we would conclude that the C variant is *associated* with the trait found in the test group.





We can do an *association study* using a small number of *candidate SNPs* chosen on the basis of a specific biological hypothesis. In the case of synaesthesia, for example, we might want to test if SNPs in a gene involved in neural connectivity are associated with synaesthesia. To do this, we would need DNA from around 200 synaesthetes.

Because no-one is certain about which kinds of genes might be involved in synaesthesia, we would also like to do a *genome-wide association study* or *GWAS*, in which thousands of SNPs from across the entire genome are tested simultaneously for association. The results of a *GWAS* are often represented using a *Manhattan plot* like the one shown below. By looking at the functions of the genes that lie in the *association peaks*, we can gain new insights into the biological basis of synaesthesia. To perform a GWAS we would need DNA from a least 1000 synaesthetes.

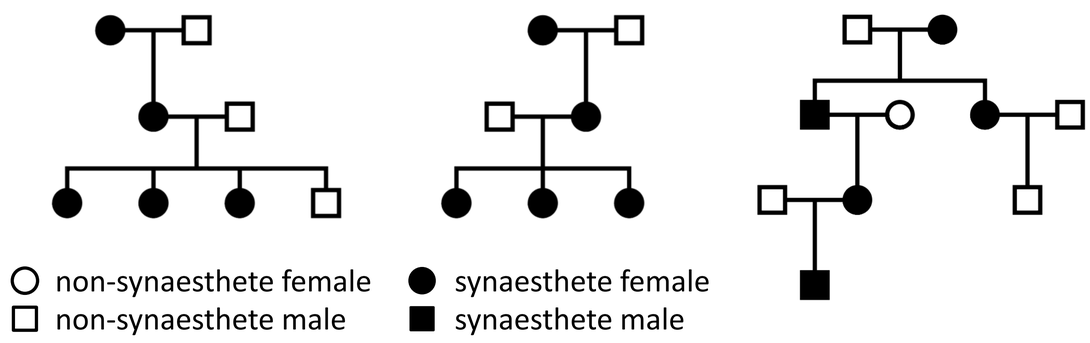


Discovering rare genetic variants: exome sequencing

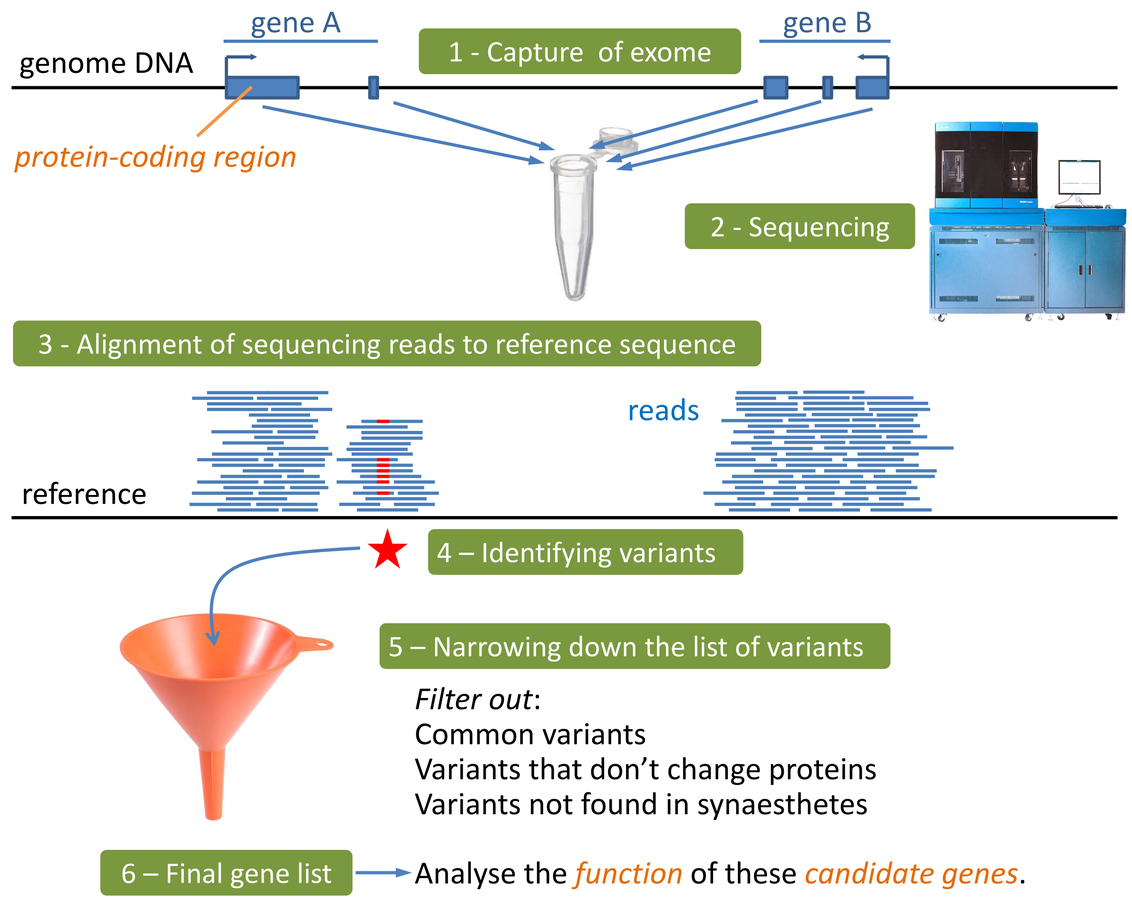
The human genome contains over 20,000 *genes*, each of which *encodes* the instructions to make a *protein* – a molecular machine with a specialized function in the body. The *protein-coding* sections of the genome are collectively referred to as the *exome*. The exome makes up less than 2% of the total genome, but it is where *rare variants* with large effects on biological traits are most likely to be found.

Using *next-generation DNA sequencing* technology, we are sequencing the exomes of individuals from families in which multiple people have synaesthesia. We then screen the sequences for rare genetic variants that are shared by related synaesthetes, and may therefore contribute to synaesthesia in these families. By looking at the function of the genes affected by these rare variants, we can learn about the molecular processes that underlie synaesthesia.

Example families for exome sequencing



Exome sequencing workflow



Perspectives

Association studies and *exome sequencing* provide two complementary approaches to identify genes that may be involved in synaesthesia. Findings from these two approaches will be integrated to assess whether they converge on the same biological pathways.

The identification of specific *genes* and *gene networks* involved in synaesthesia will for the first time shed light on the molecular basis of this intriguing neurodevelopmental condition. In addition, it may also have wider implications for neurobiology, contributing to our understanding of how neural connectivity is established in the developing brain.

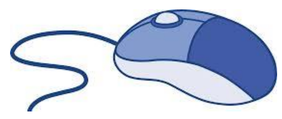
Take part in our research!

To enable us to investigate the genetic basis of synaesthesia we need:

- **Large numbers of unrelated synaesthetes** for association studies.
- **Families with several synaesthetes and non-synaesthetes** for exome sequencing.

Researchers: We are interested to hear from researchers who work with synaesthetes who may like to participate in our genetic studies.
Synaesthetes: Join our study! You do not have to travel to take part. We will assess your synaesthetic experiences online and provide a saliva kit by post to obtain your DNA (this is completely painless!).

Contact: synaesthesia@mpi.nl
Visit www.mpi.nl/synaesthesia for information



Acknowledgements

We thank Arianna Vino for assistance with DNA sample preparation, and the Human Genetics Department at Radboud University Medical Centre for performing DNA sequencing.