The genetic basis of dyslexia: The \textit{KIAA0319} gene

Silvia Paracchini

Wellcome Trust Centre for Human Genetics
University of Oxford
Summary

• Overview about the genetics of dyslexia

• Focus on the chromosome 6p locus and the KIAA0319 candidate: the gene and the protein

• The role of the dyslexia candidates in brain development
Dyslexia—definition

- Specific difficulty in learning to read that cannot be explained by deficits in intelligence, learning opportunity or any evident neurological or sensory handicap.

- Reading ability is a continuous measure; there is no an universally accepted threshold to classify an individual for being affected.
Reading

Two component processes:

- **Phonological Processing**
  - Breakdown of words into their constituent phonemes or speech sounds through the use of a set of pronunciation rules
    - AUTOMATIC “Ah-toe-Mah-tik”

- **Orthographic Processing**
  - Holistic recognition of words based on the memorised spatial appearance of letters. Requires previous word exposure
    - YACHT
Dyslexia - genetic component

- Risk in population: 5-10%
- Male / female: 3/2
- Risk in sibling of affected individuals: 38-43%
- Twin studies
  - Concordance rate: 68% in MZ versus 38% in DZ
- Complex trait, influenced by both environmental and multiple genetic factors (quantitative trait loci = QTL)
Theories of dyslexia

The biological and cognitive causes underlying the development of dyslexia are not clear.

Several theories have been proposed:

– Phonological deficit
– Auditory processing
– Visual processing
– Motor control
– Magnocellular theory

We expect the dyslexia susceptibility genes to be expressed in the brain but we don’t have a functional model.
Overview of genetic analysis results

- Regions identified by linkage analysis that might contain QTLs for dyslexia susceptibility

- Candidate genes identified by association analysis or translocation breakpoint refinement
Our Dyslexia Samples

- Genetic samples from 264 nuclear UK families:
  - Divided in sample 1 (89 families) and sample 2 (175 families)
  - All contain at least one dyslexic child (scoring on single-word reading test at least 2 SD below what predicted by test of verbal and non-verbal reasoning).
  - At least 68% contain another child with reading-related problems.
  - Total of 1153 individuals:
    - including 630 siblings measured for reading and language related quantitative traits.

- Genetic samples from 155 twin-based nuclear US families from the CLDRC:
  - Families selected for having at least one member with documented reading difficulty.
  - Total of 675 individuals:
    - including 365 siblings measured for reading and language related quantitative traits.
Our sample - Quantitative phenotype

- Essentially 6 core traits tested for:
  - Orthographic coding: Irregular words recognition - \textit{OC-irreg}
    - Example: Yacht
  - Phonological Decoding: Non-word recognition - \textit{PD}
    - Example: Siglop
  - Orthographic coding: Forced Choice - \textit{OC-choice}
    - Example: Rain versus rane
  - Word reading - \textit{READ}
  - Spelling ability - \textit{SPELL}
  - Phonological Awareness - \textit{PA} (Spoonerisms)
    - Example: lazy dog -> daisy log

- High correlation between measures: 0.3-0.8
Chromosomme 6

linkage analysis in 89 UK families, 224 siblings - SAMPLE 1

Effect of IQ on linkage analysis

6 quantitative reading-related measures

- PD
- OC-irreg
- OC-choise
- Spell
- Read
- PA

Francks et al. AJHG 2004
Initial SNP association analysis

\[ P = 0.0004 \]
OC-irreg IQ adjusted

Statistical analysis by \textbf{QTDT} - Test of association for quantitative traits in nuclear families
Association analysis and LD evaluation in 89 UK families

Association becomes stronger in families selected for severity

P < 0.00001

LD map
Association
QTDT
LD blocks
Genes
ALDH5A1
KIAA0319
TTRAP
THEM2
FLJ2619

SNP map
25 kb
### Association P-values in the UK and US most severe cases

A study in a completely independent sample detected association with 2 SNPs located in block B (AJHG Apr 2005)

| Marker       | LD region | Risk allele | OC-irreg | OC-choice | PD   | READ | SPELL | PA   | Risk allele | OC-choice | PD   | READ | SPELL | PA   |
|--------------|-----------|-------------|----------|-----------|------|------|--------|------|-------------|-----------|------|------|--------|------|------|------|------|--------|-----------|------|------|--------|------|------|------|------|
| rs699463     | A         | 1           |          |           | 0.0032 |      |        |      | 0.0279     | 0.0153    |      |      |        | 0.012 |
| rs4504469    | B         | 1           | 0.0011   | 0.0082    | 0.004 |      |        |      | 0.0101     | 0.0157    |      |      |        |      |      |
| rs2179515    | B         | 1           | 0.0012   | 0.0131    | 0.0064 | 0.0232|        |      | 0.0036     | 0.0157    |      |      |        |      |      |
| rs761101     | B         | 1           | 0.0025   | 0.0057    | 0.0066 | 0.0325|        |      | 0.0007     | 0.0155    |      |      |        |      |      |
| rs6456624    | B         | 1           |          | 0.0005    | 0.0045 | 0.0033| 0.0157|      | 0.0017     | 0.0155    |      |      |        |      |      |
| rs2328846    | B         | 1           |          | 0.0007    | 0.0017 | 0.0033| 0.0155|      | 0.0001     | 0.0155    |      |      |        |      |      |
| rs2235676    | B         | 2           | 0.023    | 0.0009    | 0.0041 |      |        |      | 0.0002     | 0.0155    |      |      |        |      |      |
| rs2038137    | B         | 1           | 0.0013   | 0.0026    | 0.0002 | 0.0061|        |      | 0.0032     | 0.0155    |      |      |        |      |      |
| k_pr_del     | B         | 1           | 0.0011   | 0.0032    | 0.0002 | 0.0086|        |      | 0.0006     | 0.0155    |      |      |        |      |      |
| k_pr_1       | B         | 2           | 0.0006   | 0.0003    | 0.0373 | 0.0003| 0.0116|      | 0.0017     | 0.0155    |      |      |        |      |      |
| rs1555090    | B         | 1           | 0.001    | 0.0029    | 0.0003 | 0.0131|        |      | 0.0006     | 0.0155    |      |      |        |      |      |
| rs3033236    | B         | 2           | 0.0134   | 0.0104    | 0.0073 |      |        |      | 0.0006     | 0.0155    |      |      |        |      |      |
| rs2143340    | B         | 2           | 0.01     | 0.0003    | 0.0115 |      |        |      | 0.0006     | 0.0155    |      |      |        |      |      |
| rs1061925    | B         | 2           | 0.0009   | 0.0005    | 0.0009 |      |        |      | 0.0006     | 0.0155    |      |      |        |      |      |
| tt_th_del    | C         | 1           |          |          | 0.0132 |      |        |      | 0.015      | 0.0415    |      |      |        |      |      |
| rs926529     | C         | 1           |          |          | 0.0132 |      |        |      | 0.015      | 0.0415    |      |      |        |      |      |
| rs1885211    | C         | 1           |          |          | 0.0132 |      |        |      | 0.015      | 0.0415    |      |      |        |      |      |
| th_ex_3      | C         | 2           | 0.0332   |          |      |      |        |      | 0.0415    | 0.0415    |      |      |        |      |      |
| rs3756814    | C         | 2           |          |          | 0.0332 |      |        |      | 0.015      | 0.0415    |      |      |        |      |      |

**Colour-coded P-values:**
- **0.01 < P < 0.05**
- **0.001 < P < 0.01**
- **P < 0.001**
Haplotype analysis

Block B

Tagging SNPs

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>UK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4504469</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>(1) 1</td>
<td>1</td>
<td>41</td>
</tr>
<tr>
<td>(2) 2</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>(3) 1</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>(4) 2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>(5) 1</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Tagging SNPs

rs2038137

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>UK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(2) 2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(3) 1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(4) 2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(5) 1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Tagging SNPs

rs2143340

<table>
<thead>
<tr>
<th>Haplotypes</th>
<th>UK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(2) 2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>(3) 1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(4) 2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(5) 1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

UK

% P(read)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

US

% P(read)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

MUTATION SCREENING DID NOT REVEAL ANY OBVIOUS DISRUPTIVE MUTATION
Allele-specific quantitative gene expression assay

Select cell lines heterozygous for the risk haplotype

Measure of relative quantitative differences in gene expression using:

Coding SNPs

\[ \begin{align*}
\text{C} & \quad \text{G} \\
\text{T} & \quad \text{A}
\end{align*} \]

SNPs at promoter site

\[ \begin{align*}
\text{T} & \quad \text{A}
\end{align*} \]

Make cDNA

PCR

Primer extension

Mass spectrometry

Measure of peak areas is proportional to relative abundance of the 2 alleles

Immunoprecipitate chromatin (HaploChIP)

PCR
HaploChIP principle

SNP    RNA Pol II binding site

Protein-DNA crosslinking *in vivo*

Protein digestion

PCR and allele quantification

Immunoprecipitation with antibody specific to RNA Pol II

THE RELATIVE QUANTITY OF THE 2 ALLELES IS A MEASURE OF RNA POL II AFFINITY TO THE 2 HAPLOTYPES

Nat Genet 2003, 33:469-75
Risk haplotype effect on gene expression

3 CEPH lymphoblastoid cell lines:
- Heterozygous for risk haplotype
- Heterozygous for SNPs within the transcripts or the promoters of the three genes

Paracchini et al, HMG 2006
The risk haplotype is associated to a reduced expression of the KIAA0319 gene

Risk v non-risk
\[ P = 2 \times 10^{-7} \]

Risk v non-risk
\[ P = 7 \times 10^{-16} \]
The DCDC2 gene

Two studies found association within DCDC2, less than 200kb away from KIAA0319:

• **Meng et al., 2005**: Association with 2 SNPs + identification of an intronic deletion somehow linked to dyslexia in 153 US families from CLDR

• **Schumacher et al., 2006**: Two-SNPs haplotype within intron seven associated in two independent German trio samples. Association is stronger in severe sub-groups
### KIAA0319 v DCDC2

**Oxford/Cardiff collaboration**

Harold et al, Mol Psychiatry 2006

<table>
<thead>
<tr>
<th>Marker</th>
<th>Gene</th>
<th>Location</th>
<th>Risk allele&lt;sup&gt;a&lt;/sup&gt;</th>
<th>OC-irreg PD</th>
<th>OC-choice</th>
<th>READ</th>
<th>SPELL</th>
<th>PA</th>
<th>Risk allele&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs793862_rs807701</td>
<td>DCDC2</td>
<td>Intron 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs793862</td>
<td>DCDC2</td>
<td>Intron 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs807701</td>
<td>DCDC2</td>
<td>Intron 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs807724</td>
<td>DCDC2</td>
<td>Intron 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCDC2_deletion</td>
<td>DCDC2</td>
<td>Intron 3</td>
<td>2</td>
<td>0.0298</td>
<td>0.0428</td>
<td>0.0478</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1087266</td>
<td>DCDC2</td>
<td>Intron 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Selected sibships from the Oxford sample

n = 313 siblings, 126 families

Cardiff cases and controls

n = 350 / n = 273

Harold et al, Mol Psychiatry 2006
Chromosome 6p: result summary

Paracchini et al, ARGG in press
Chromosome 6p: result interpretation

- Associations are different signals of a unique causal mutations
- Different association are the results of different ascertainment and analysis criteria.
- Each gene contribute to a specific sub-group of dyslexia
- Definitive answer would come from the identification of the causal genetic variants
## The Colorado case

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of families</th>
<th>Selection criteria</th>
<th>Associated gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deffenbacher et al. (2004)</td>
<td>114</td>
<td>A sib with severe score in at least one trait</td>
<td>DCDC2 KIAA0319</td>
</tr>
<tr>
<td>Francks et al. (2004)</td>
<td>126</td>
<td>A sib with severe score on a composite measure of traits contributing to the linkage</td>
<td>KIAA0319</td>
</tr>
<tr>
<td>Meng et al. (2005)</td>
<td>153</td>
<td>No selection</td>
<td>DCDC2</td>
</tr>
</tbody>
</table>

> Each gene contribute to specific subgroup of dyslexia

> The analysis is very sensitive to sample selection
KIAA0319: In situ expression

Mouse brain E13.5

KIAA0319

Positive control

Negative control

Mouse brain E15.5

KIAA0319

Positive control

Paracchini et al, HMG 2006
KIAA0319: In situ expression

Early fetal human brain
In utero RNA interference

Joe LoTurco, University of Connecticut

Glial-guided neuronal migration

Cortical plate
Intermediate zone
Ventricular zone

THEM2
TTRAP

KIAA0319
KIAA0319 - rescue

Paracchini et al, HMG 2006
Neuronal migration and dyslexia

- Description of different dyslexic brains revealed cortical malformations mainly in the frontal region and in the left language areas consistent with abnormal neuronal migration (Galaburda’s work):
  - Ectopias (small neuronal congregations in an abnormal layer location)
  - Dysplasia (loss of cortical neurons organisation)

- The neuronal migration disorder of periventricular nodular heterotopia has been found to be associated with an impairment in reading skills in presence of otherwise normal intelligence (Chang et al., 2005).
Neuronal migration and KIAA0319

- KIAA0319 is a transmembrane protein exposing **PKD domains** outside the cell.

- PKD domains have **cell adhesion** properties.

- It is possible that KIAA0319 is required for appropriate cell adhesion between **migrating neurons** and the glial fibers during the development of the neocortex.
## Other candidate genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Independent replications</th>
<th>Functional variants</th>
<th>Functional mechanism</th>
<th>Brain expression</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYX1C1</td>
<td>?</td>
<td>NO</td>
<td></td>
<td>YES</td>
<td>Taipale (2003)</td>
</tr>
<tr>
<td>ROBO1</td>
<td>None</td>
<td>NO</td>
<td>Gene expression</td>
<td>YES</td>
<td>Hannula-Jouppi (2005)</td>
</tr>
<tr>
<td>MRPL19</td>
<td>2 (same study)</td>
<td>NO</td>
<td>Gene expression</td>
<td>YES</td>
<td>Anthoni (2007)</td>
</tr>
<tr>
<td>C2ORF3</td>
<td>2 (same study)</td>
<td>NO</td>
<td>Gene expression</td>
<td>YES</td>
<td>Anthoni (2007)</td>
</tr>
</tbody>
</table>
# DYX1C1: replication attempts

<table>
<thead>
<tr>
<th>Reference</th>
<th>Proband's disorder</th>
<th>Country of origin</th>
<th>Most significantly reported P-values$^a$ for individual SNPs or haplotypes within <em>DYX1C1</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Taipale et al. (95)</td>
<td>Dyslexia</td>
<td>Finland</td>
<td>-3G &gt; A: 0.002 (A) 1249G &gt; T: 0.006 (T) -3G &gt; A:1249G &gt; T: 0.015 (A:T)</td>
</tr>
<tr>
<td>Scerri et al. (81)$^b$</td>
<td>Dyslexia</td>
<td>U.K.</td>
<td>n/s 1249G &gt; T: 0.0076 (G) -3G &gt; A:1249G &gt; T: 0.0140 (G:G) 0.0182 (G:T)</td>
</tr>
<tr>
<td>Wigg et al. (105)</td>
<td>Dyslexia</td>
<td>Canada$^c$</td>
<td>-3G &gt; A: 0.021 (G) n/s -3G &gt; A:1249G &gt; T: 0.026 (G:G)</td>
</tr>
<tr>
<td>Cope et al. (17)</td>
<td>Dyslexia</td>
<td>U.K.</td>
<td>n/s n/s n/s</td>
</tr>
<tr>
<td>Marino et al. (62)</td>
<td>Dyslexia</td>
<td>Italy</td>
<td>n/s n/s n/s</td>
</tr>
<tr>
<td>Meng et al. (65)</td>
<td>Dyslexia</td>
<td>U.S.$^c$</td>
<td>n/s n/s n/t</td>
</tr>
<tr>
<td>Bellini et al. (4)</td>
<td>Dyslexia</td>
<td>Italy</td>
<td>n/s n/s n/t</td>
</tr>
<tr>
<td>Ylisaukko-Oja et al. (110)</td>
<td>Autism</td>
<td>Finland</td>
<td>n/s n/s n/s</td>
</tr>
<tr>
<td>Wigg et al. (106)</td>
<td>ADHD</td>
<td>Canada$^c$</td>
<td>n/s n/s n/t</td>
</tr>
</tbody>
</table>

*Paracchini et al, ARGG in press*
Dyslexia candidates and brain development

- KIAA0319, DCDC2 and DYX1C1 have been implicated in neuronal migration

- ROBO1 is a receptor for the chemorepellent SLIT. The SLIT-ROBO system controls axon branching, commissural axon pathfinding and neuronal migration.
The million dollar question

How can neuronal migration genes specifically affect reading skills?

Are these genes specifically expressed in reading-related cortical regions?
Brain expression profile

Paracchini et al. ARGG in press
Brain expression profile

Paracchini et al, ARGG in press
EXPRESSION PROFILES IN DIFFERENT TISSUES

DCDC2

ROBO1

KIAA0319

DYX1C1

Tissue

- BMR: Bone marrow
- SPL: Spleen
- TMS: Thymus
- BRN: Brain
- SPC: Spinal cord
- HRT: Heart
- MSL: Skeletal muscle
- LVR: Liver
- PNC: Pancreas
- PST: Prostate
- KDN: Kidney
- LNG: Lung

GeneNote – expression arrays

Normalized intensity

- DCDC2
- ROBO1
- KIAA0319
- DYX1C1
Genes and reading skills

- The candidates are not expressed in reading-specific cortical area. They are also expressed in tissues different from the brain.

- **WE DON'T EXPECT TO FIND THE GENE FOR READING** (as FOXP2 in not the gene for language!!)

- Suboptimal neuronal migration may result in cortical abnormalities that affect reading-related regions. The same genes may also affect other cognitive functions.

- Cortical abnormalities in specific regions would depend on multiple gene-gene, gene-environment interactions.
The NeuroDys Project

- Multidisciplinary project grouping 13 research groups from 10 European countries with different expertise

- Access to ~4000 samples

- Major goals:
  - Identify the dyslexia susceptibility genetic variants
  - Link genetic background to sub-groups of dyslexic phenotypes
  - Link genetic background to specific neurological markers
Conclusion

- The KIAA0319 gene is a strong candidate for dyslexia susceptibility, supported by association data in at least three independent samples.

- A specific haplotype associated to dyslexia is also associated to reduced expression of the KIAA0319 gene.

- The KIAA0319 is required for neuronal migration during the development of the neocortex.

- The other dyslexia candidates are also involved in cortex development.

- Genetics is playing a crucial role in uncovering the causes of dyslexia.
Acknowledgements

Genetic analysis

University of Oxford
Clyde Francks
Tom Scerri
Laurence MacPhie
Simon Fisher
Angela Marlow
Janet Walter
Alex Richardson
Lon Cardon
John Stein
Anthony Monaco

Colorado Study
Shelley Smith
Bruce Pennington
Richard Olson
John DeFries

Functional analysis

University of Oxford
Antonio Velayos
Brendan Keating
Julian Knight
Claudio Toma
Mēgan Dennis
Jerome Nicod
Tara Caffrey
Jennifer Taylor
Richard Wade-Martins
Anthony Monaco

Joe LoTurco
Thomas Ankur
Marugan Paramasivan
Yu Wang

Andy Copp
Sandra Castro
Cecilia Lai