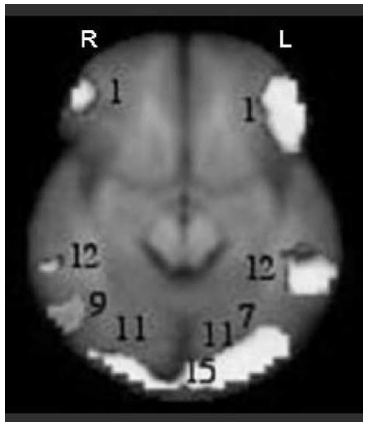
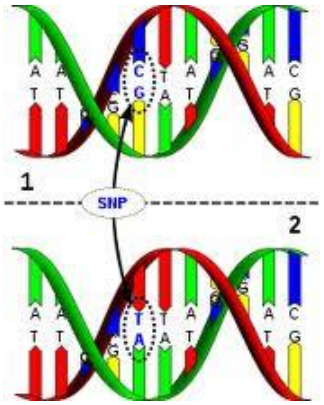
Examining the Genetic Underpinnings of Commonly Comorbid Language Disorders: Dyslexia and Language Impairment





John Eicher
Golden Helix Webinar
Department of Genetics
Yale University
May 13, 2014

Two Common Language Disorders

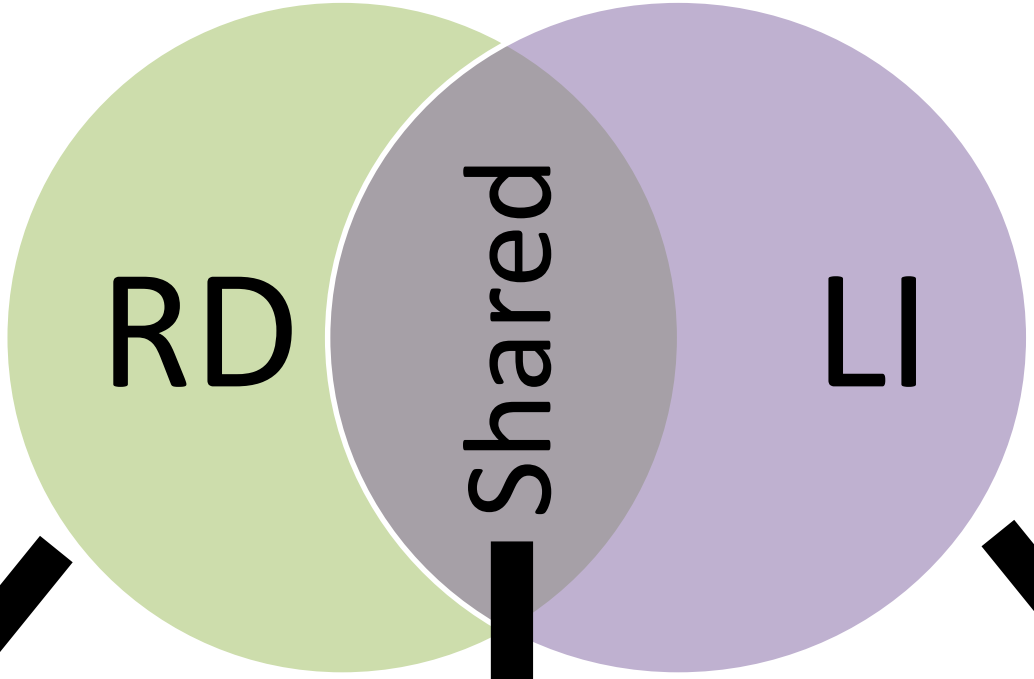
- Dyslexia/Reading Disability (RD)

Defense ≠ Di---fens ≠

- Language Impairment (LI)

 ≠ Bull---dog ≠

What are Reading Disability (RD) and Language Impairment (LI)?

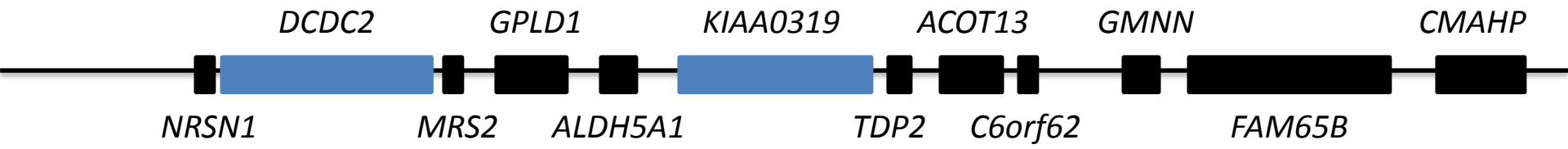


Reading Disability (RD	Comorbidity of RD and LI	airment (LI)
Prevalence: 5-17%	50% of LI cases develop RD	8%
Reading Decoding/Co	RD cases more likely to have/had LI	ehension
Phonological Processi	Phonological Impairments	ceptive Language
Written Language	Involve overall language deficits	ge

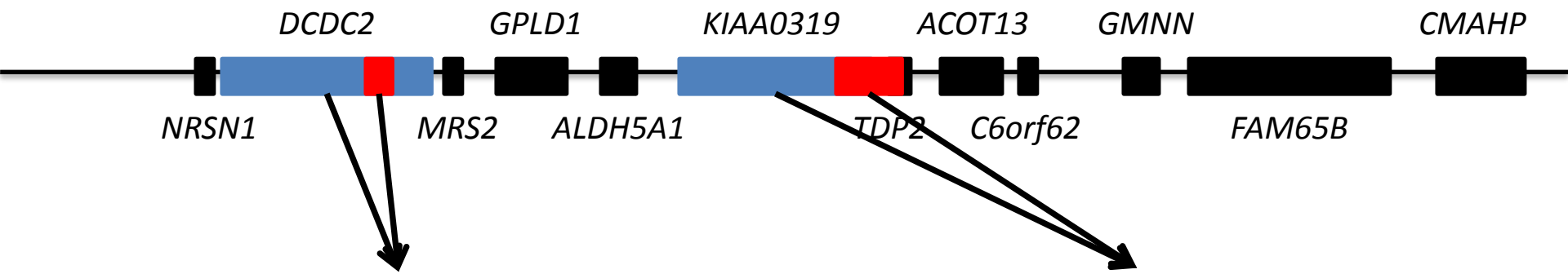
Brief History of Genetics of RD/LI

- Genetic components of RD and LI
  - Heritability estimates of RD: 54-85%
  - Heritability estimates of LI : 45-73%
- Strongest candidate genes include:
  - *DCDC2* and *KIAA0319* in DYX2 (chr. 6)

DYX2 Locus on 6p22



DYX2 Locus (Chromosome 6p22)



DCDC2

- Neuronal Migration
- Grey/white matter
- Microtubule binding domain
- Regulates expression

KIAA0319

- Neuronal Migration
- Grey/white matter
- Signaling protein
- Associated with reduced expression of *KIAA0319*

READ1

- "Regulatory Element Associated with Dyslexia 1"
- Highly polymorphic
- Replicated multiple times
- Specifically binds TF ETV6

KIAA0319 haplotype

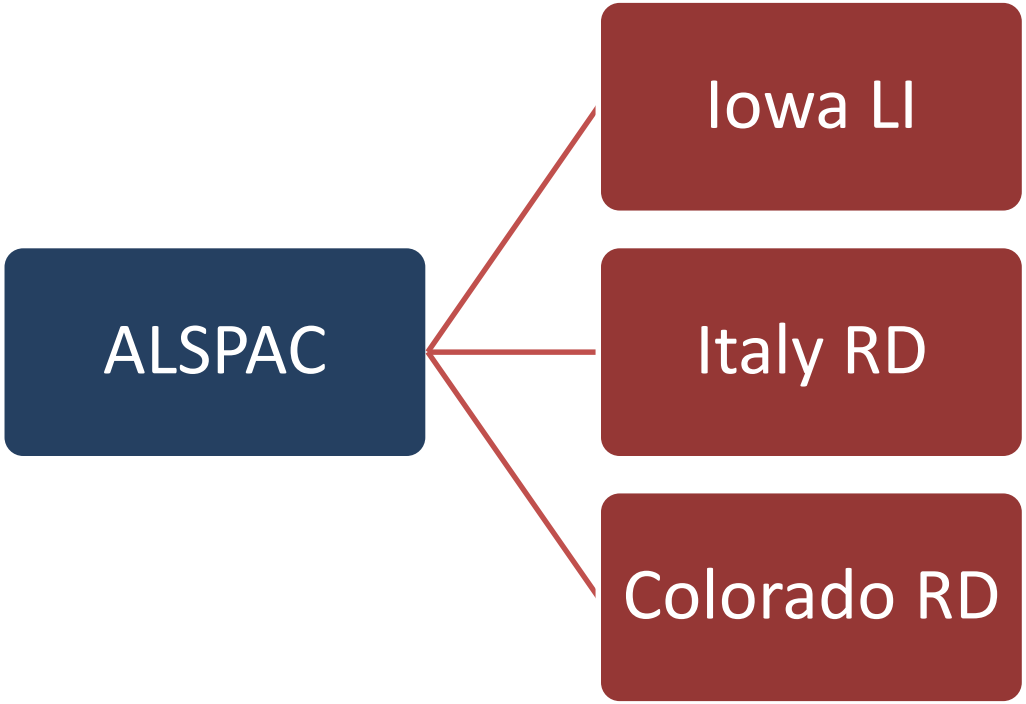
- Located within the *KIAA0319* promoter
- Overlaps *TDP2*
- Replicated multiple times
- Locus associated with RD and LI

Brief History of Genetics of RD/LI

- Genetic components of RD and LI
  - Heritability estimates of RD: 54-85%
  - Heritability estimates of LI : 45-73%
- Strongest candidate genes include:
  - *KIAA0319* and *DCDC2* in DYX2 (chr. 6)
  - *DYX1C1* in DYX1 (chr. 15)
  - *FOXP2* and *CNTNAP2* (chr. 7)
- Only one GWAS examining quantitative performance on reading and language tasks

Objectives

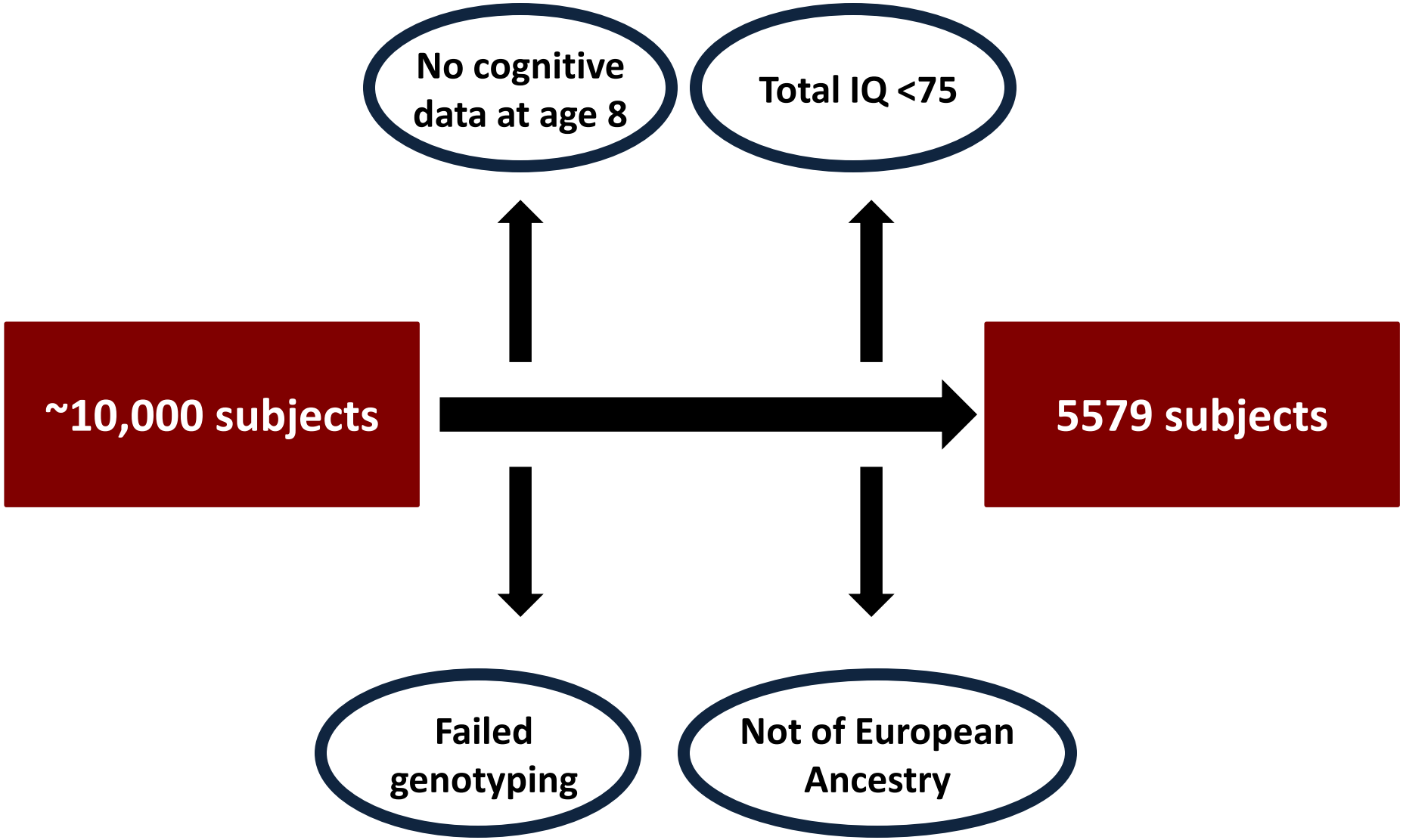
- To characterize the relationship of the DYX2 locus with RD, LI, and IQ
  - Discovery: Avon Longitudinal Study of Parents of Children (ALSPAC)
  - Replication: Iowa LI, Italy RD, Colorado RD



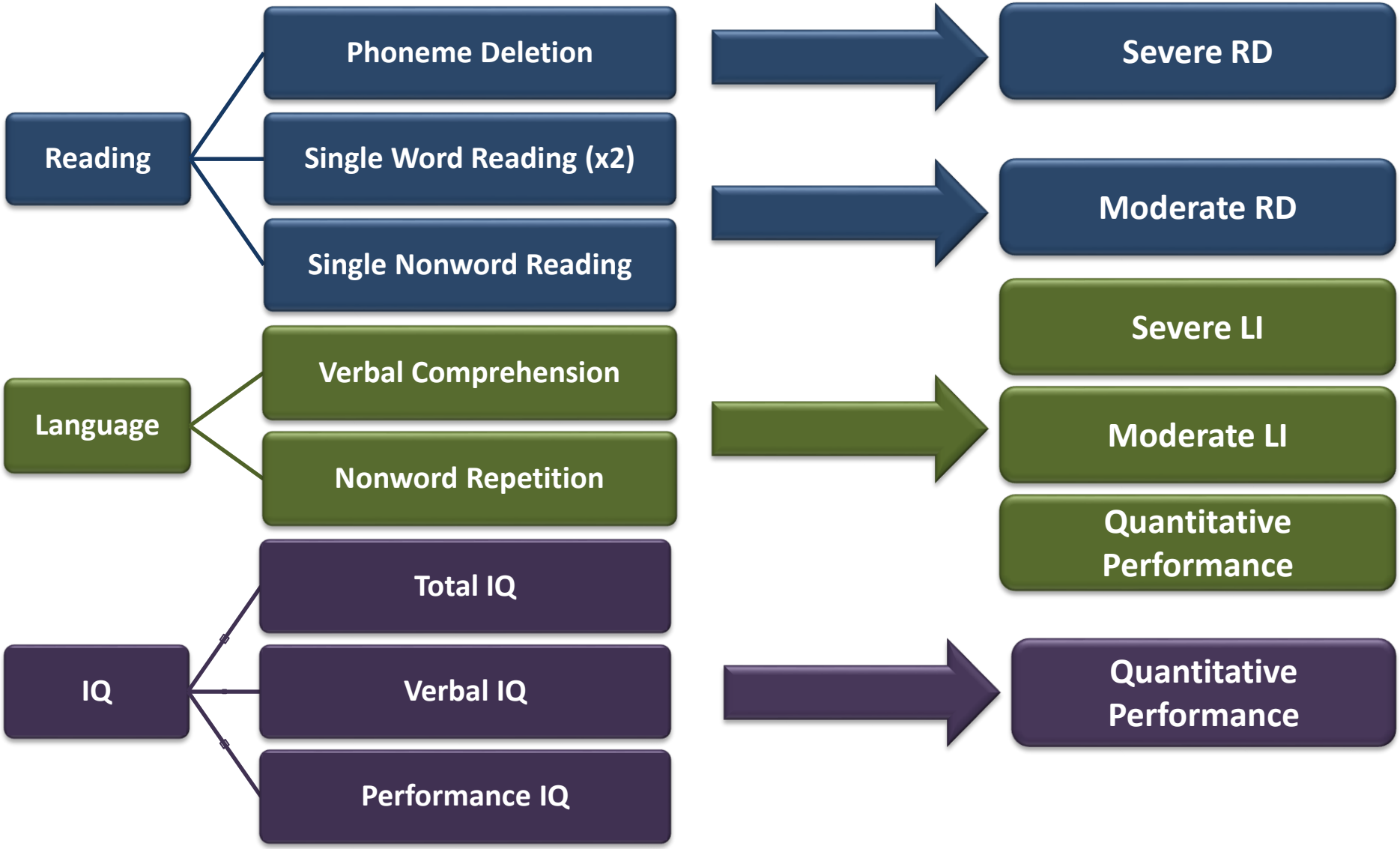
What is ALSPAC?

- Longitudinal birth cohort in Avon, UK
  - Over 10,000 pregnant women enrolled
  - Data collected on children from prenatal period to present time (approximately 21 years old)
  - Conducted at the University of Bristol
- Myriad of environmental and clinical data
  - Written language and reading
  - Verbal language and speech
  - Other neurocognitive and communicative data

Avon Longitudinal Study of Children and Parents (ALSPAC)



Phenotypes Collected in ALSPAC



Genotyping Strategy

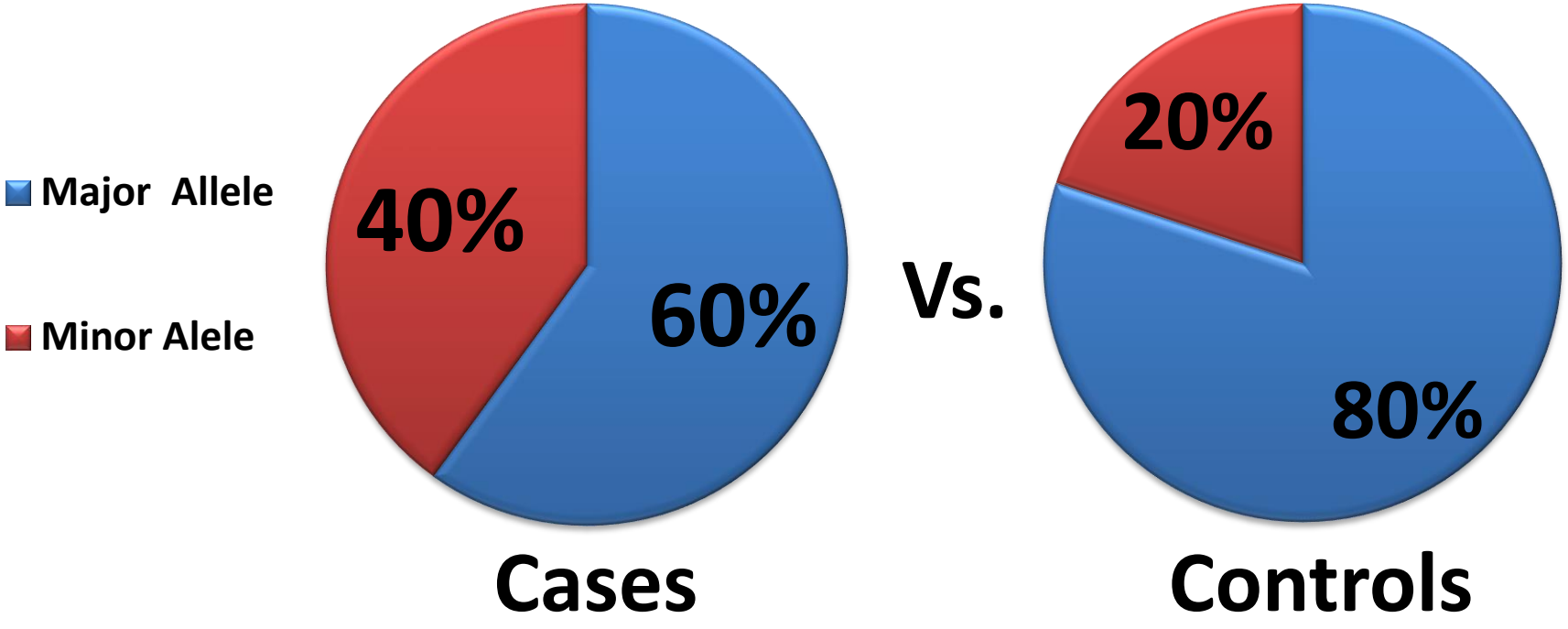
- Tag single nucleotide polymorphisms (SNPs) to capture as much variation in the DYX2 locus as possible

DYX2 Locus on 6p22

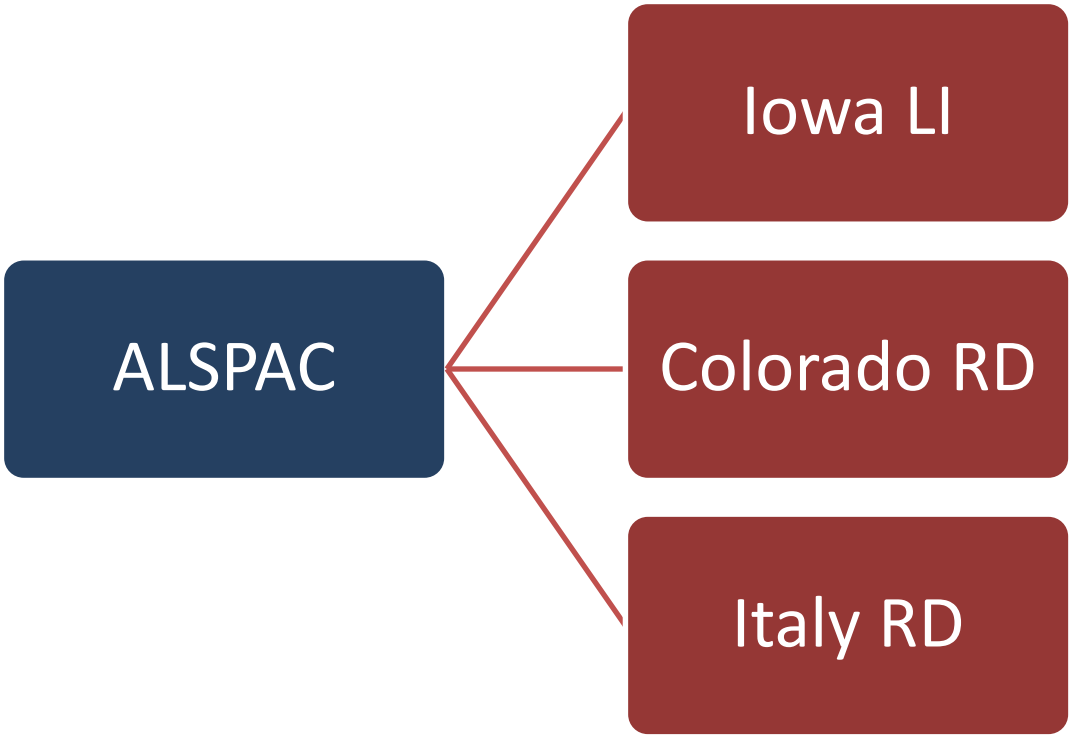
- Total of 195 markers covering ~1.4 Mb
- Completed using Sequenom MassARRAY
- Allow for unbiased association scan of locus

Association Methods

- Single Marker Analysis in SNP & Variation Suite (SVS) v7.6.4
  - Compare allele frequencies in cases and controls
  - Regress quantitative performance on genotype
- Haplotypes (Haploview v4.2) associations completed with PLINK v1.07



DYX2 Association Strategy

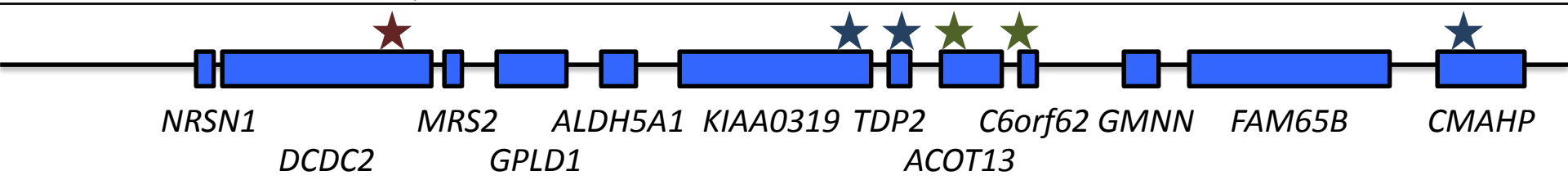


	Iowa LI	**Colorado RD**	**Italy RD**
Cohort-type	Case-control	Family-based	Family-based
Number of Subjects	428	1188	878
Number of Families	N/A	292	304
Analysis	SVS	TDT (PLINK)	TDT (PLINK)
Conditioned on:	Case-Control Status	Case-Control Status, Discriminant Score	Case-Control Status

Association with RD and LI

Phenotype	Marker	Gene	BP	Model	Odds Ratio	P-value
Severe LI	rs807694	*DCDC2*	24303383	Additive	1.8	5.70×10^{-4}
Severe LI	rs807694	*DCDC2*	24303383	Dominant	1.9	6.20×10^{-4}
Severe RD	rs10456309	*KIAA0319*	24589562	Recessive	10.5	2.00×10^{-4}
Severe RD	rs2294691	*TDP2*	24652843	Additive	1.9	5.30×10^{-4}
Severe RD	rs2294691	*TDP2*	24652843	Dominant	2.3	1.80×10^{-4}
Moderate LI	rs3777663	*ACOT13*	24700235	Additive	0.6	3.90×10^{-4}
Moderate LI	rs3756814	*C6orf62*	24705835	Additive	0.7	3.90×10^{-4}
Moderate RD	rs1562422	*CMAHP*	25044577	Dominant	1.7	8.10×10^{-4}

Phenotype	Markers	Haplotype	Gene	BP	OR	P-value
Severe RD	rs33914824, rs807694, rs707864, rs10456301, rs16889066, rs9379651	CGCGAG	*DCDC2*	24302046-24314900	3.20	6.07×10^{-5}
Severe LI	rs33914824, rs807694, rs707864, rs10456301, rs16889066, rs9379651	GACGAG	*DCDC2*	24302046-24314900	1.91	2.84×10^{-4}



Associations with Quantitative Language

Marker	Gene	BP	Model	Slope	P-value
rs9295626	*KIAA0319*	24587339	Additive	0.064	7.30×10^{-4}
rs9348646	*FAM65B*	24820219	Additive	-0.129	2.60×10^{-4}

Markers	Haplotype	Gene	BP	Slope	P-value
rs2817201, rs9295626	AT	*KIAA0319*	24585214-24587339	0.064	7.40×10^{-4}
rs10456309, rs4576240, rs17307478, rs9356939, rs7763790, rs6456621	GGTCAC	*KIAA0319*	24589562-24618511	0.064	5.90×10^{-4}
rs6935076, rs2038137, rs3756821, rs1883593, rs3212236	AGATA	*KIAA0319*	24639223-24648455	0.078	8.70×10^{-5}

Associations with IQ

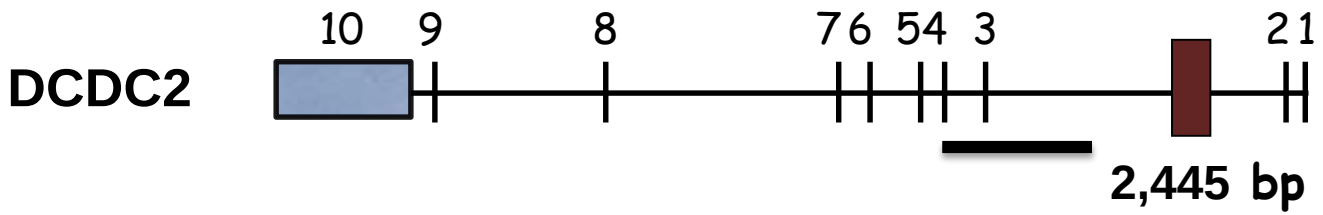
Phenotype	Marker	Gene	BP	Model	Slope	P-value
Total IQ	rs2328791	N/A	23736848	Additive	-1.18	7.50×10^{-4}
Total IQ	rs2328791	N/A	23736848	Recessive	-3.36	4.20×10^{-4}
Verbal IQ	rs9295626	*KIAA0319*	24587339	Additive	1.39	4.30×10^{-4}
Verbal IQ	rs7763790	*KIAA0319*	24615063	Additive	-1.38	4.80×10^{-4}
Verbal IQ	rs6935076	*KIAA0319*	24644322	Additive	1.15	5.20×10^{-4}
Verbal IQ	rs9348646	*FAM65B*	24052526	Additive	-1.14	6.60×10^{-4}

Markers	Haplotype	Gene	BP	Slope	P-value
rs2817201, rs9295626	AT	*KIAA0319*	24585214-24587339	1.42	3.78×10^{-4}
rs10456309, rs4576240, rs17307478, rs9356939, rs7763790, rs6456621	GGTCAC	*KIAA0319*	24589562-24618511	-1.40	5.69×10^{-4}
rs6456624, rs6935076, rs2038137, rs3756821, rs1883593, rs3212236	AGATA	*KIAA0319*	24639223-24648455	1.81	1.45×10^{-5}
rs3777663, rs3756814, rs6931809, rs6916186, rs6933328, rs17491647	TGTGGA	*ACOT13/ C6orf62*	24700235-24713723	-1.56	7.42×10^{-4}

Replication of ALSPAC DYX2 Results

Marker	Gene	Iowa LI Case Control		Italy RD Case Control		Colorado RD Case Control		Colorado RD Discriminant Score	
		OR	p	OR	p	OR	p	Slope	p
rs33914824	*DCDC2*	**2.2**	**0.034**	0.9	0.768	1.1	0.847	0.023	0.934
rs807694	*DCDC2*	**1.9**	**0.028**	0.9	0.786	0.9	0.853	-0.025	0.919
rs707864	*DCDC2*	**1.6**	**0.017**	1.0	0.840	1.2	0.446	-0.246	0.101
rs9295626	*KIAA0319*	1.1	0.579	**0.6**	**0.0055**	1.0	0.823	-0.158	0.169
rs10456309	*KIAA0319*	0.5	0.073	0.7	0.189	0.4	0.206	**0.628**	**0.0133**
rs4576240	*KIAA0319*	1.1	0.825	**1.9**	**0.0027**	1.1	0.862	-0.052	0.754
rs9356939	*KIAA0319*	**4.0**	**0.018**	0.8	0.069	1.3	0.151	-0.116	0.254
rs6456621	*KIAA0319*	**2.2**	**0.019**	1.6	0.405	1.8	0.366	-0.458	0.104
rs1883593	*KIAA0319*	1.3	0.169	**1.6**	**0.0052**	1.3	0.239	-0.108	0.395
rs3777663	*ACOT13*	**0.7**	**0.016**	**0.6**	**0.0052**	1.0	0.908	0.101	0.345
rs3756814	*C6orf62*	**0.7**	**0.005**	**0.7**	**0.023**	0.9	0.600	-0.003	0.980
rs6931809	*C6orf62*	**1.4**	**0.023**	**1.4**	**0.017**	1.2	0.491	-0.096	0.382
rs6933328	*C6orf62*	0.9	0.612	0.9	0.613	1.0	0.827	**0.215**	**0.0515**
rs9348646	*FAM65B*	0.9	0.358	1.1	0.535	1.4	0.144	**-0.415**	**0.00051**
rs1562422	*CMAHP*	1.0	0.793	1.0	0.796	**0.6**	**0.093**	-0.030	0.840

DCDC2 Risk Haplotype and READ1



.....(GAGAGGAAGGAAA)2....(GGAA)8....DelGAAA.....(GGAA)2...(GGAA)3/4...(GGAA)2.....

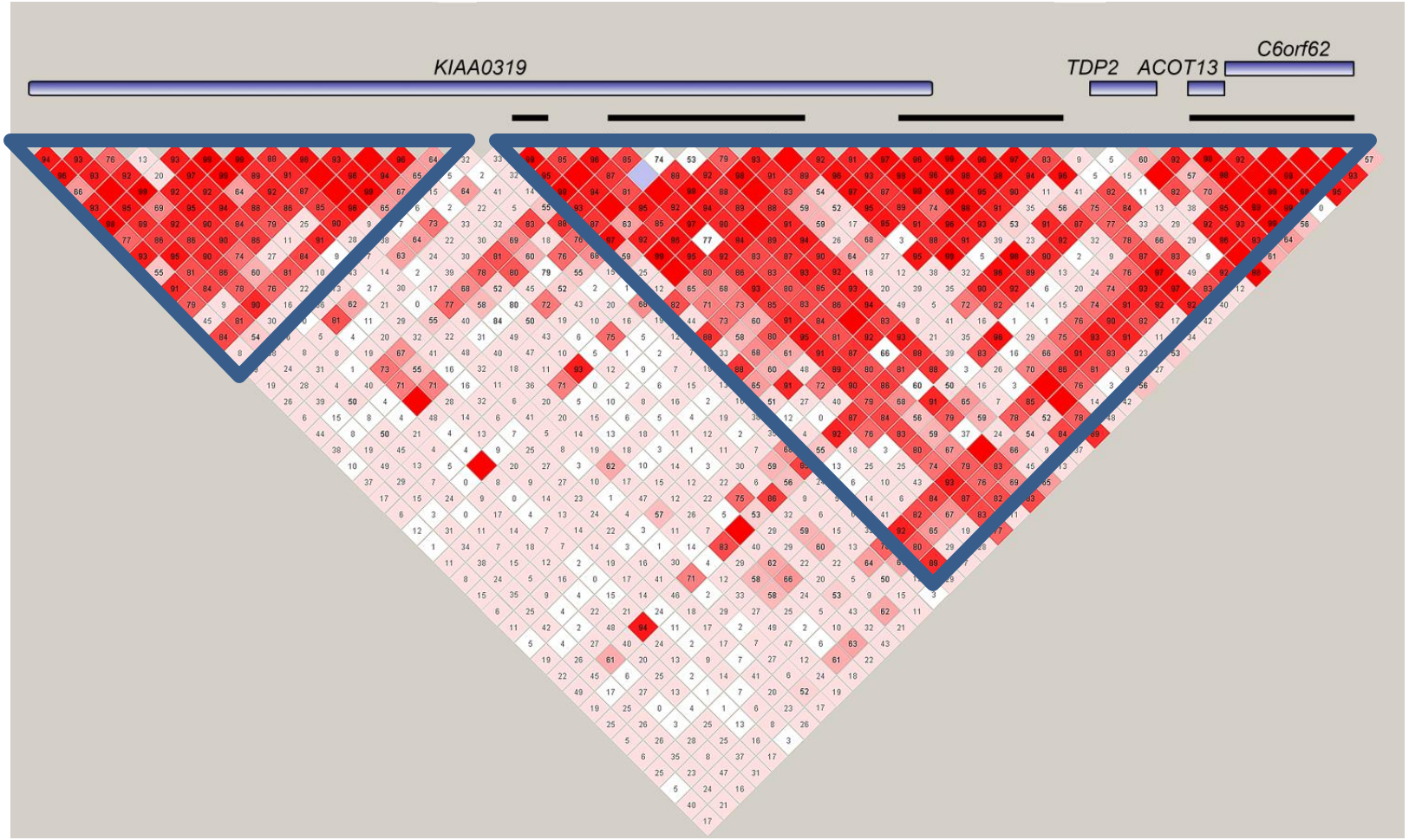
Association Data				
Haplotype	**Phenotype**	**Haplotype Freq.**	**Odds Ratio**	**P-value**
CGCGAG	Severe RD	0.0236	3.20	6.07×10^{-5}
GACGAG	Severe LI	0.0364	1.91	2.84×10^{-4}

Powers et al. AJHG 93(1), 19-28

Association of READ1 to RD and LI

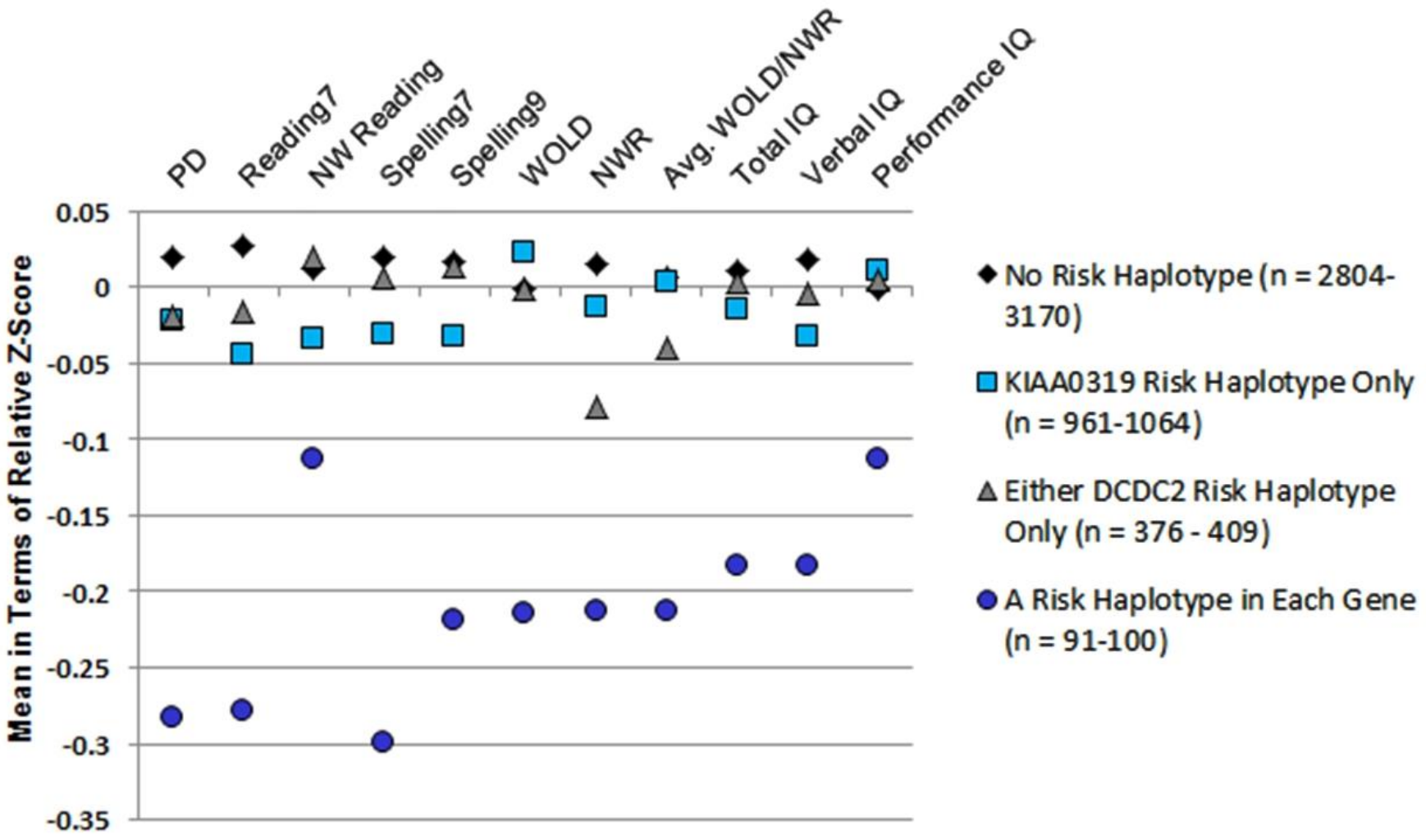
	Severe LI		Severe RD	
Allele/Grouping	**OR**	**P-value**	**OR**	**P-value**
Allele 3	0.77	0.255	0.575	0.179
Allele 4	0.78	0.141	1.28	0.239
Allele 5	0.84	0.488	2.37	5.80×10^{-5}
Allele 6	1.65	5.95×10^{-3}	1.53	0.010
Allele 10	0.90	0.603	0.919	0.795
Clade 1	1.73	7.40×10^{-5}	1.89	6.20×10^{-5}
Longer Alleles	1.68	8.96×10^{-3}	2.22	1.17×10^{-3}
Shorter Alleles	0.80	0.292	0.506	0.096

Linkage Disequilibrium in *KIAA0319*



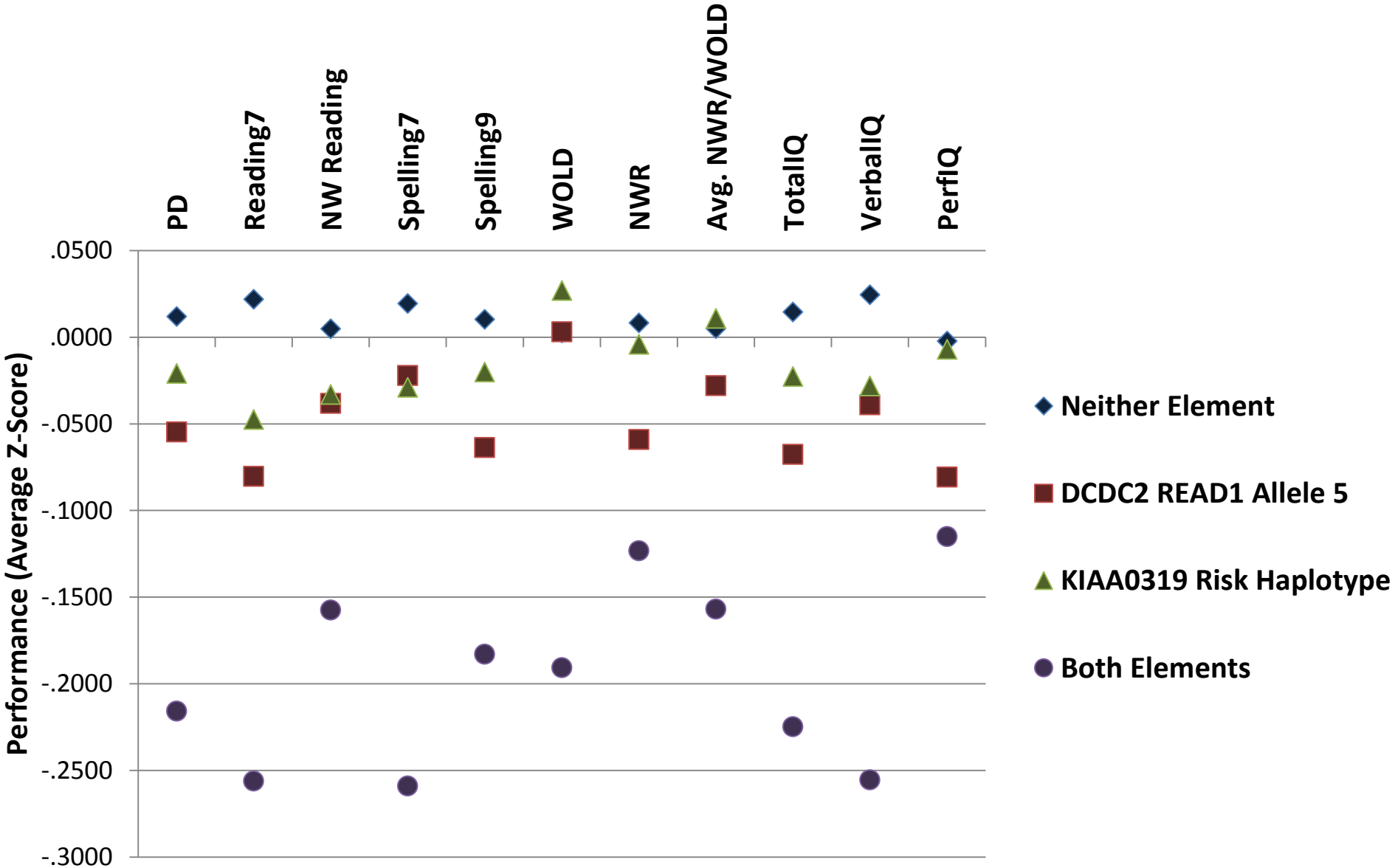
- Two LD Blocks: (1) 3’ Half of *KIAA0319* and (2) 5’ Half of *KIAA0319/TDP2/ACOT13/C6orf62*

Interaction between *DCDC2* and *KIAA0319* Haplotypes



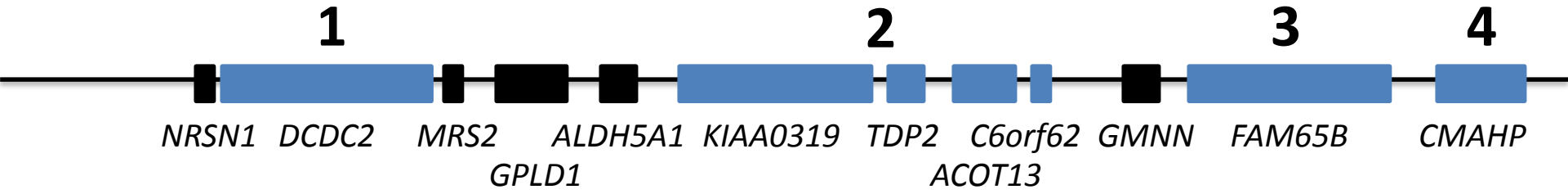
Powers et al. AJHG 93(1), 19-28

Interaction between READ1 Allele 5 and *KIAA0319* Haplotype



Summary of DYX2 Results

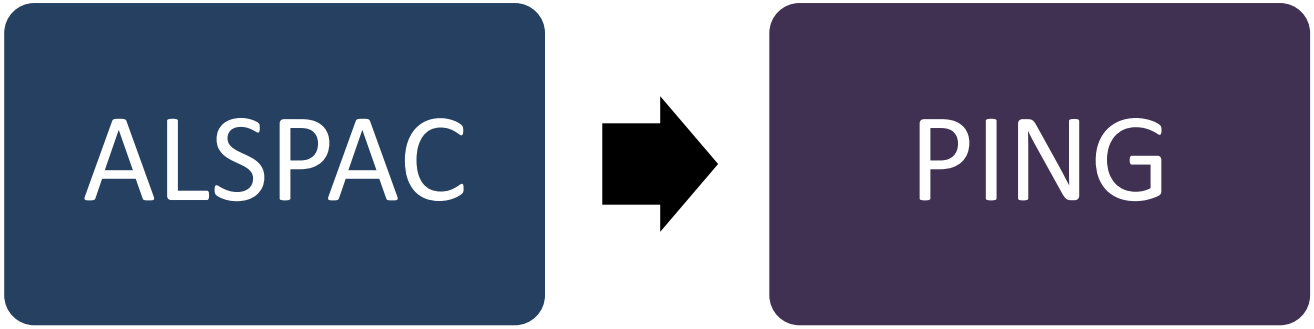
- 4 DYX2 loci associated with RD, LI, and/or IQ
  1. *DCDC2*: READ1 element
  2. *KIAA0319*: KIAA0319 risk haplotype (5' region)
  3. *FAM65B*
  4. *CMAHP*
- Evidence for interaction between READ1 in *DCDC2* and the *KIAA0319* risk haplotype



But what about new genes?

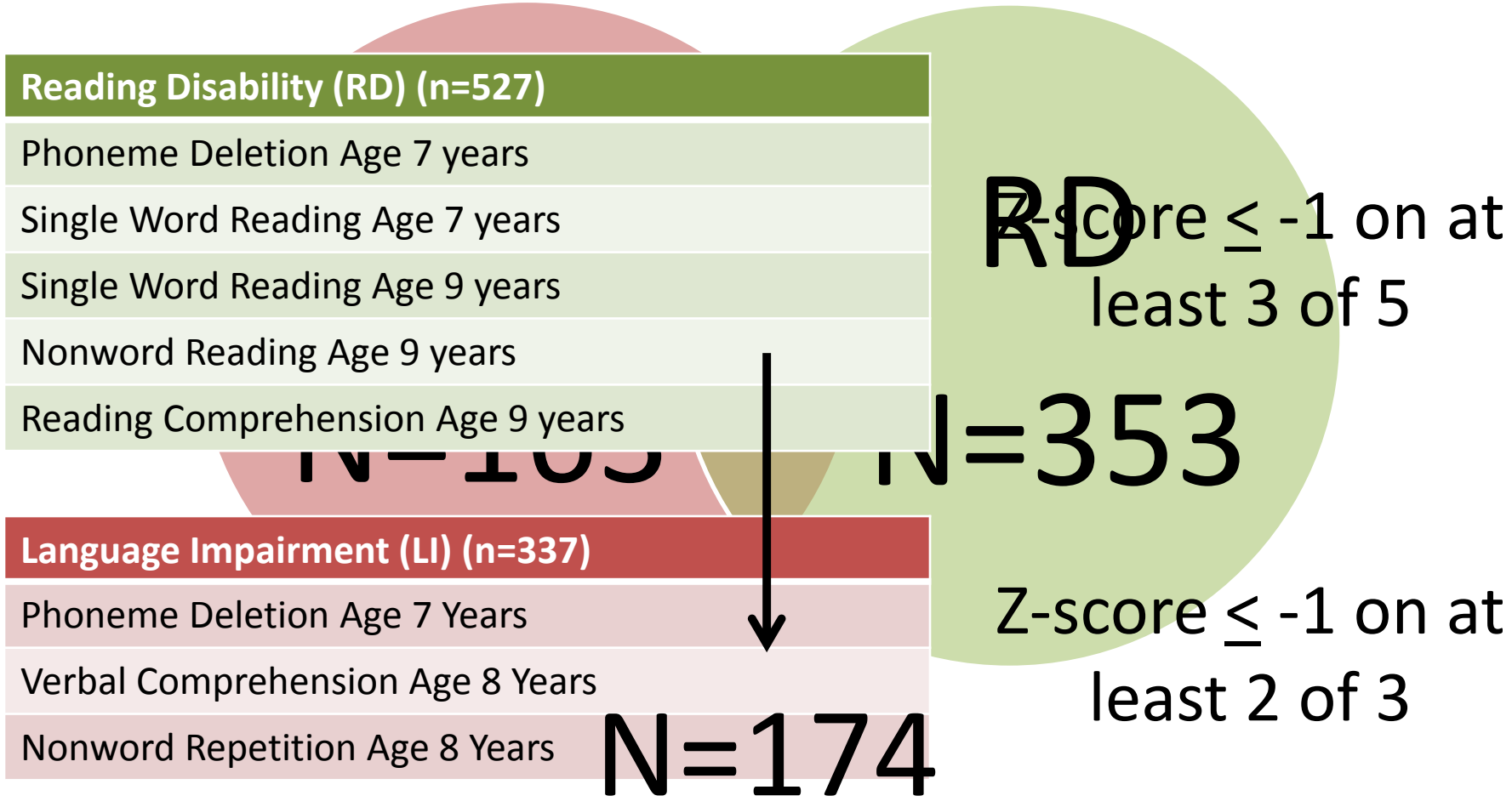
- So far, I characterized the relationship of a known risk locus with RD, LI, and IQ
- These analyses do not implicate novel regions
- Hypothesis-free methods, including genome-wide association (GWAS) and whole exome sequencing (WES), can identify novel risk genes

GWAS of Comorbid RD and LI



- GWAS, scanning entire genome with ~500,000 SNPs, comparing allele frequencies in cases and controls
- Here, cases are define as subjects with both RD and LI in ALSPAC
- Top 10 associated markers were moved forward for replication in PING

RD and LI Case Definitions in ALSPAC



Analytical Strategy

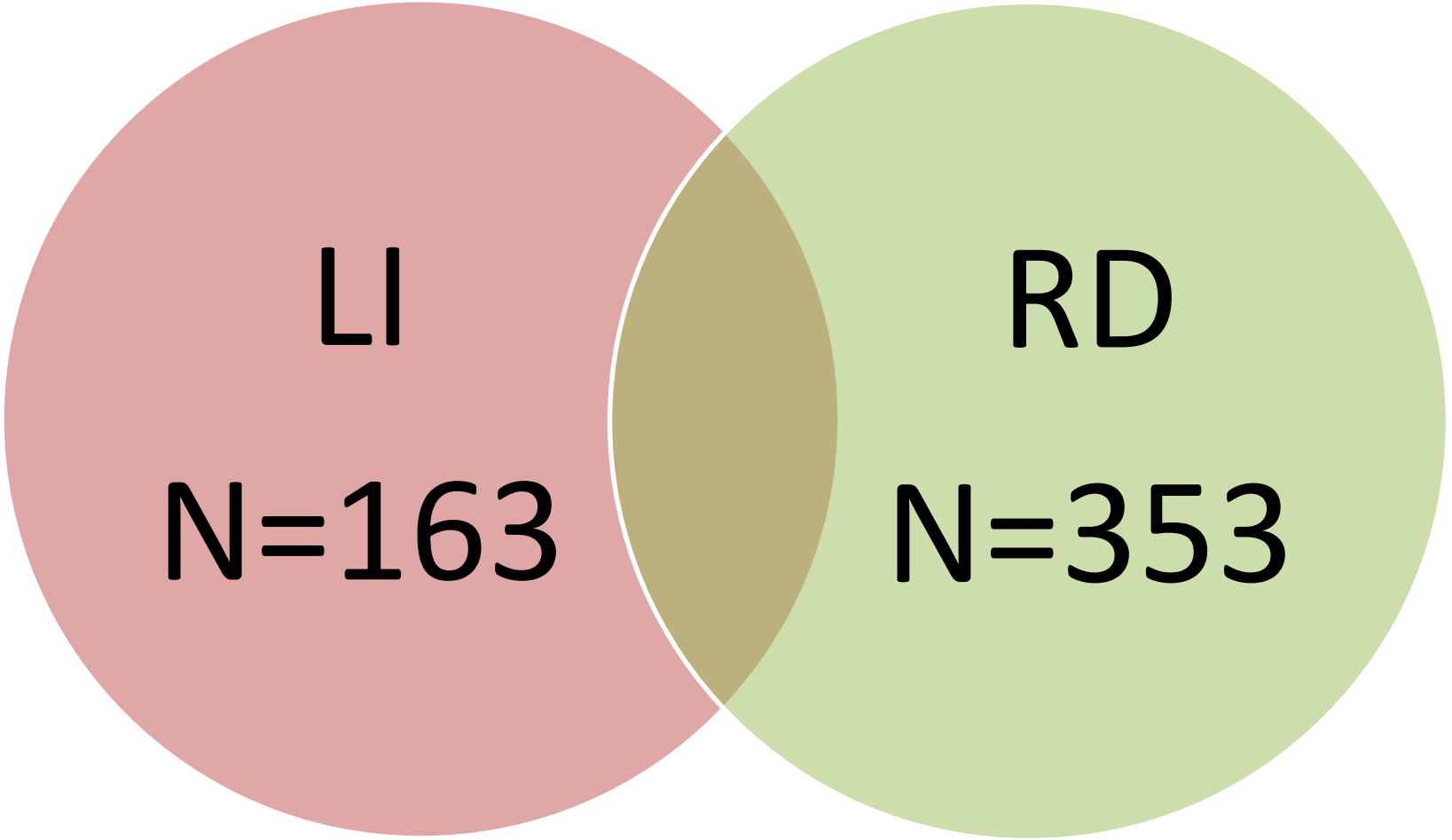
- GWAS analyses ($\alpha = 1.00 \times 10^{-7} = 0.05 / 500{,}000$)
  - First, performed in cases with both RD and LI (174 cases, 4117 controls)
  - Second, performed in cases with: (1) only RD (n=353) and (2) only LI (n=163)
- All associations performed in individuals of European descent

Top 10 Associations for RD and LI

Marker	Chr	Base Pair	Minor Allele	MAF Aff	MAF Unaff	Gene	OR	P-value
rs12636438	3	22038281	G	0.3017	0.1927	*ZNF385D*	1.811	5.45×10^{-7}
rs1679255	3	22022938	C	0.3006	0.1923	*ZNF385D*	1.805	6.87×10^{-7}
rs9521789	13	109917621	C	0.5201	0.3879	*COL4A2*	1.710	7.59×10^{-7}
rs1983931	13	109916103	G	0.5201	0.3896	*COL4A2*	1.698	1.06×10^{-6}
rs9814232	3	21948179	A	0.2931	0.1886	*ZNF385D*	1.784	1.30×10^{-6}
rs7995158	13	109909718	A	0.5201	0.3911		1.687	1.44×10^{-6}
rs6573225	14	58354640	C	0.1965	0.1122		1.935	1.56×10^{-6}
rs4082518	10	17103032	T	0.3103	0.2049	*CUBN*	1.746	2.17×10^{-6}
rs442555	14	58365937	C	0.1983	0.1149		1.905	2.38×10^{-6}
rs259521	3	21942154	T	0.2902	0.1885	*ZNF385D*	1.761	2.42×10^{-6}

Eicher et al. (2013) Genes Brain Behav 12(8), 792-801

Associations with Non-Comorbid Cases



Top 10 Associations for RD only

Marker	Chr	Base Pair	Minor Allele	MAF Aff	MAF Unaff	Gene	Odds Ratio	P-value
rs180950	10	115697957	G	0.456	0.369		1.431	5.16×10^{-6}
rs2590673	8	126037337	G	0.133	0.083		1.697	5.85×10^{-6}
rs892100	19	50772522	C	0.228	0.162	*OPA3*	1.526	6.92×10^{-6}
rs1792745	18	51955991	T	0.187	0.129		1.558	1.22×10^{-5}
rs12546767	8	126151747	C	0.152	0.099	*KIAA0196*	1.618	1.32×10^{-5}
rs12634033	3	146524529	C	0.135	0.087		1.646	1.80×10^{-5}
rs892270	12	105002956	G	0.534	0.451	*NUAK1*	1.395	2.16×10^{-5}
rs10887149	10	124156994	A	0.278	0.357	*PLEKHA1*	0.069	2.25×10^{-5}
rs10041417	5	33218502	T	0.226	0.164		1.489	2.58×10^{-5}
rs6792971	3	68468217	C	0.111	0.068	*FAM19A1*	1.703	2.59×10^{-5}

Eicher et al. (2013) Genes Brain Behav 12(8), 792-801

Top 10 Associations for LI only

Marker	Chr	Base Pair	Minor Allele	MAF Aff	MAF Unaff	Gene	Odds Ratio	P-value
rs482700	4	116286939	G	0.3896	0.2588	*NDST4*	1.827	1.40×10^{-7}
rs7695228	4	116309516	T	0.3920	0.2636	*NDST4*	1.801	2.94×10^{-7}
rs1940309	4	116306410	T	0.3865	0.2606	*NDST4*	1.788	4.14×10^{-7}
rs505277	4	116248257	T	0.3773	0.2528	*NDST4*	1.791	4.35×10^{-7}
rs476739	4	116248997	A	0.3773	0.2529	*NDST4*	1.79	4.41×10^{-7}
rs867036	4	116381578	C	0.3957	0.2696	*NDST4*	1.774	5.31×10^{-7}
rs867035	4	116381423	C	0.3957	0.2697	*NDST4*	1.773	5.45×10^{-7}
rs2071674	4	2366882	T	0.0920	0.0389	*ZFYVE28*	2.503	1.90×10^{-6}
rs7694946	4	116413588	C	0.3620	0.2526	*NDST4*	1.678	8.95×10^{-6}
rs4823324	22	44616787	C	0.2914	0.4143	*ATXN10*	0.581	9.30×10^{-6}

Eicher et al. (2013) Genes Brain Behav 12(8), 792-801

Pediatric Imaging Neurocognition Genetics (PING) Study

- 1300 typically developing children from 10 sites across the United States
  - Primary Coordinating Site at UCSD
  - Yale University one of the recruiting sites
- Each individual underwent:
  - Neurocognitive assessments, including oral reading and receptive language tasks
  - Neuroimaging battery (structural and DTI MRI)
  - Genetic information (GWAS)

Replication of ALSPAC Markers in PING

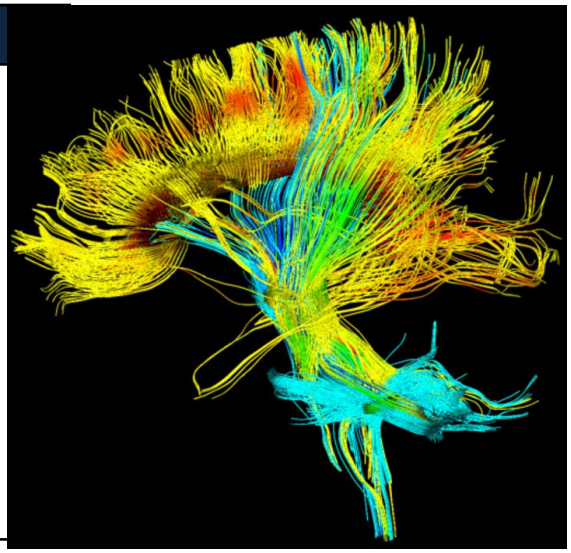
Marker	Allele	MAF	Gene	Oral Reading Test		Picture Vocabulary Test	
				Beta	P-value	Beta	P-value
rs12636438	**G**	**0.161**	***ZNF385D***	**-0.1867**	**0.9452**	**-2.88**	**0.004173***
rs1679255	**G**	**0.292**	***ZNF385D***	**-1.84**	**0.5016**	**-3.048**	**0.002445***
rs9521789	G	0.437	*COL4A2*	-0.3411	0.7332	0.8647	0.3877
rs476739	A	0.265	*NDST4*	0.5406	0.5891	0.5159	0.6062
rs505277	A	0.280	*NDST4*	0.5406	0.5891	-0.3452	0.7301
rs482700	G	0.278	*NDST4*	0.5498	0.5828	-0.05341	0.9574
rs7695228	A	0.295	*NDST4*	0.6258	0.5318	0.09991	0.9205
rs867036	G	0.378	*NDST4*	0.2605	0.7946	-0.1414	0.8876
rs867035	G	0.377	*NDST4*	0.2961	0.7673	-0.1565	0.8757
rs1940309	A	0.281	*NDST4*	0.6049	0.5456	0.1296	0.8969

Eicher et al. (2013) Genes Brain Behav 12(8), 792-801

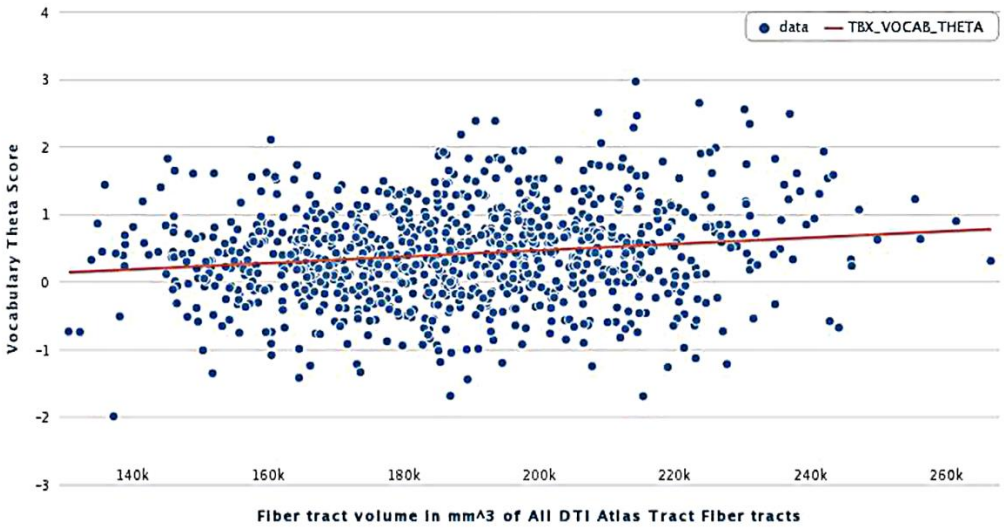
Imaging-Genetics in PING

- Examined in 332 European subjects in PING
- Covariates included in model: (1) Age, (2) Gender, (3) MRI Scanner, (4) Handedness, and (5) Socioeconomic Status
- Examined 16 fiber tracts of interest

Fiber Tract of Interest	Abbreviation
All Fiber Tracts	All
Inferior Longitudinal Fasiculus	ILF
Inferior Fronto-Occipital Fasiculus	IFO
Superior Longitudinal Fasiculus	SLF
Temporal Superior Longitudinal Fasiculus	tSLF
Parietal Superior Longitudinal Fasiculus	pSLF
Striatal Inferior Frontal Cortex	SIFC
Corpus Callosum	CC

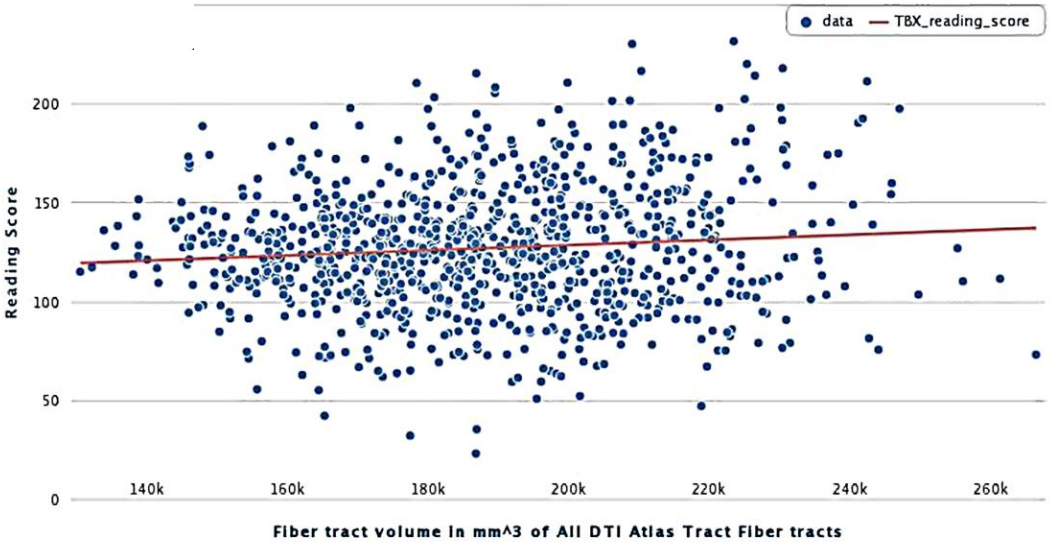


Fiber Tract Volumes Correlated with Reading and Language Performance



Increased fiber tract volume associated with increased performance on a receptive vocabulary task (p=0.000602)

Increased fiber tract volume associated with increased performance on a oral reading task (p=0.03596)

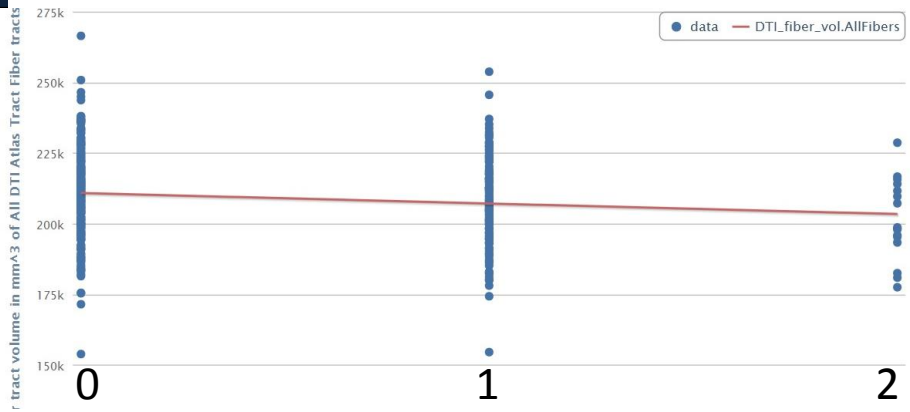


Eicher et al. (2013) Genes Brain Behav 12(8), 792-801

Association of *ZNF385D* with Overall Fiber Tract Volumes

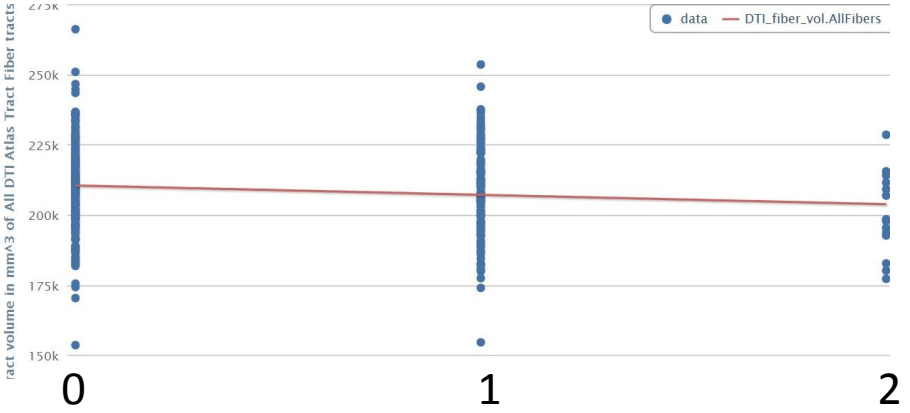
	rs1679255		rs12636438	
	Slope	**P-value**	**Slope**	**P-value**
All	-3329.9	0.044*	-3717.9	0.023*
Right All	-1731.4	0.039*	-1965	0.017*
Left All	-1616.3	0.055	-1775.6	0.033*
Right ILF	-251.3	0.011*	-234.4	0.016*
Left ILF	-256.9	0.0088**	-254.6	0.009**
Right IFO	-200.8	0.032*	-190	0.041*
Left IFO	-221	0.012*	-226.3	0.009**
Right SLF	-168.1	0.06	-206	0.02*
Left SLF	-199.5	0.022*	-212.9	0.013*
Right tSLF	-170.8	0.011*	-180.7	0.0068**
Left tSLF	-163.1	0.023*	-169.9	0.016*
Right pSLF	-153.1	0.079	-182.4	0.034*
Left pSLF	-112.2	0.18	-125.3	0.131
Right SIFC	-148.8	0.052	-165.6	0.029*
Left SIFC	-34.54	0.66	-54.3	0.48
CC	-977.1	0.15	-1181.6	0.081

Fiber Tract Vol (mm^3) vs. rs1679255



Number of copies of minor allele

Fiber Tract Vol (mm^3) vs. rs12636438



Number of copies of minor allele

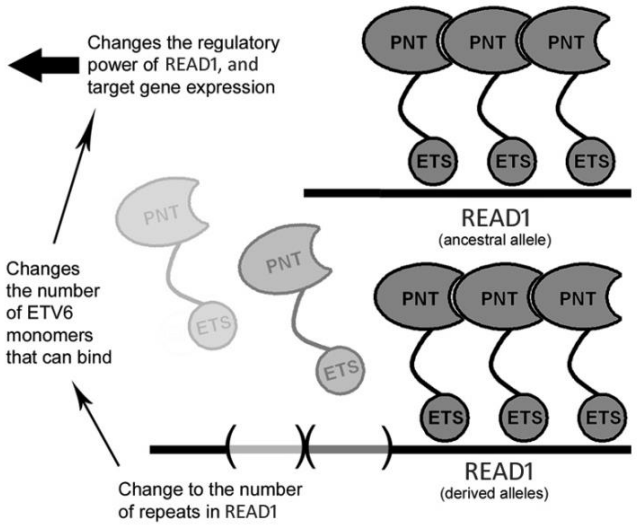
Eicher et al. (2013) Genes Brain Behav 12(8), 792-801

Summary of GWAS Findings

- *ZNF385D* associated cases with comorbid RD and LI as well as receptive vocabulary
- Two unreplicated associations
  - *NDST4* with LI
  - *COL4A2* with comorbid RD and LI
- *ZNF385D* influences overall fiber tract volumes
- Fiber tract volumes associated with reading and language performance

Discussion

- Importance of gene regulation in RD and LI
  - Rare, coding variants likely result in more severe phenotypes (e.g. gross neural abnormalities)
  - *ZNF385D* and *FOXP2* as transcription factors
  - READ1 in *DCDC2* and KIAHap in *KIAA0319*



- Preliminary evidence of biological interaction between *KIAA0319* and *DCDC2*
- Other genomic targets of READ1/ETV6 complex?

Future Work

- Meta-analyses across cohorts
  - Completed GWASes in ALSPAC and PING
  - Collaborators performed GWAS of these traits in independent cohorts
- Functional follow-up of *ZNF385D*
  - Alter *ZNF385D* expression to examine its effects on gene expression genome-wide
  - Determine where *ZNF385D* may bind across the genome (e.g. ChIP-seq)

Collaborator Acknowledgements

- All subjects and families!
- ALSPAC
  - Susan Ring
  - Laura Miller
  - George Davey-Smith
  - Beate St. Pourclain
- University of Iowa
  - J. Bruce Tomblin
  - Kathyrn Mueller
- University of Nebraska
  - Shelley Smith
- University of Denver
  - Bruce Pennington
- University of Colorado-Boulder
  - Richard Olson
  - John DeFries
  - Erik Willcutt
- Scientific Institute, IRCCS Eugenio Medea
  - Cecilia Marino
  - Sara Mascheretti
- Pediatrics Imaging Neurocognition Genetics Consortium
  - Terry Jernigan
  - Anders Dale
  - Cinnamon Bloss
  - Burcu Darst
- Keck Center at Yale University
  - Shrikant Mane
  - Irina Tikhonova
  - Anna Rogers
- Yong Kong
- Funding Sources
  - NINDS (Gruen) and NIDCD (Eicher)
  - Manton Foundation (Gruen)

Gruen Lab

- **Jeffrey R. Gruen**
- **Natalie R. Powers**
- Joan Bosson-Heenan
- Angela Montgomery
- Asya Asghar
- Katherine Loftus
- Nini Jain
- Devon Capiello
- Jocelyn Otero
- Josue Rodriguez
- Renny Parrilla