INTRODUCTION TO MOLECULAR GENETICS

Andrew McQuillin
Molecular Psychiatry Laboratory
UCL Division of Psychiatry
22 Sept 2017

Learning Objectives

Understand:

- The distinction between Quantitative Genetic and Molecular Genetic research
- The structure of DNA
- How DNA functions as the hereditary material
  - How DNA is packaged
  - How DNA function is regulated, etc
- The different types of genetic diversity in human populations
- Common terms used
- The broad types of genotype-phenotype relationships

What are the origins of individual differences in human behaviour and complex traits?

Quantitative Genetic research

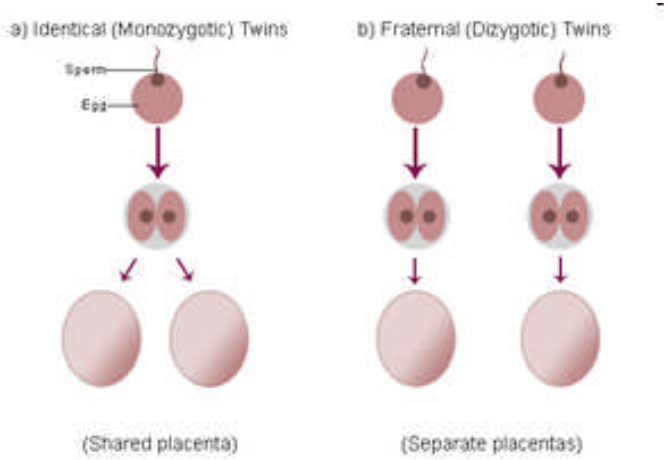
- How much do genetic (and environmental) factors contribute to individual differences in a measured trait or susceptibility to disease?

- Nature AND Nurture
  - and the correlation and interaction between the two

Why do we estimate heritability?

- Quantitative Genetic studies tell us:
- The importance of genetic and environmental influence
- Provides a best case scenario for genetic prediction
- Allows us to estimate the success of gene identification efforts

Twins - a natural experiment



Genotype 100% vs 50%
Sex
Age
Parents
In utero environment?
Early life / home environment

Molecular Genetic Research

- The study of the structure and function of genes at the molecular level

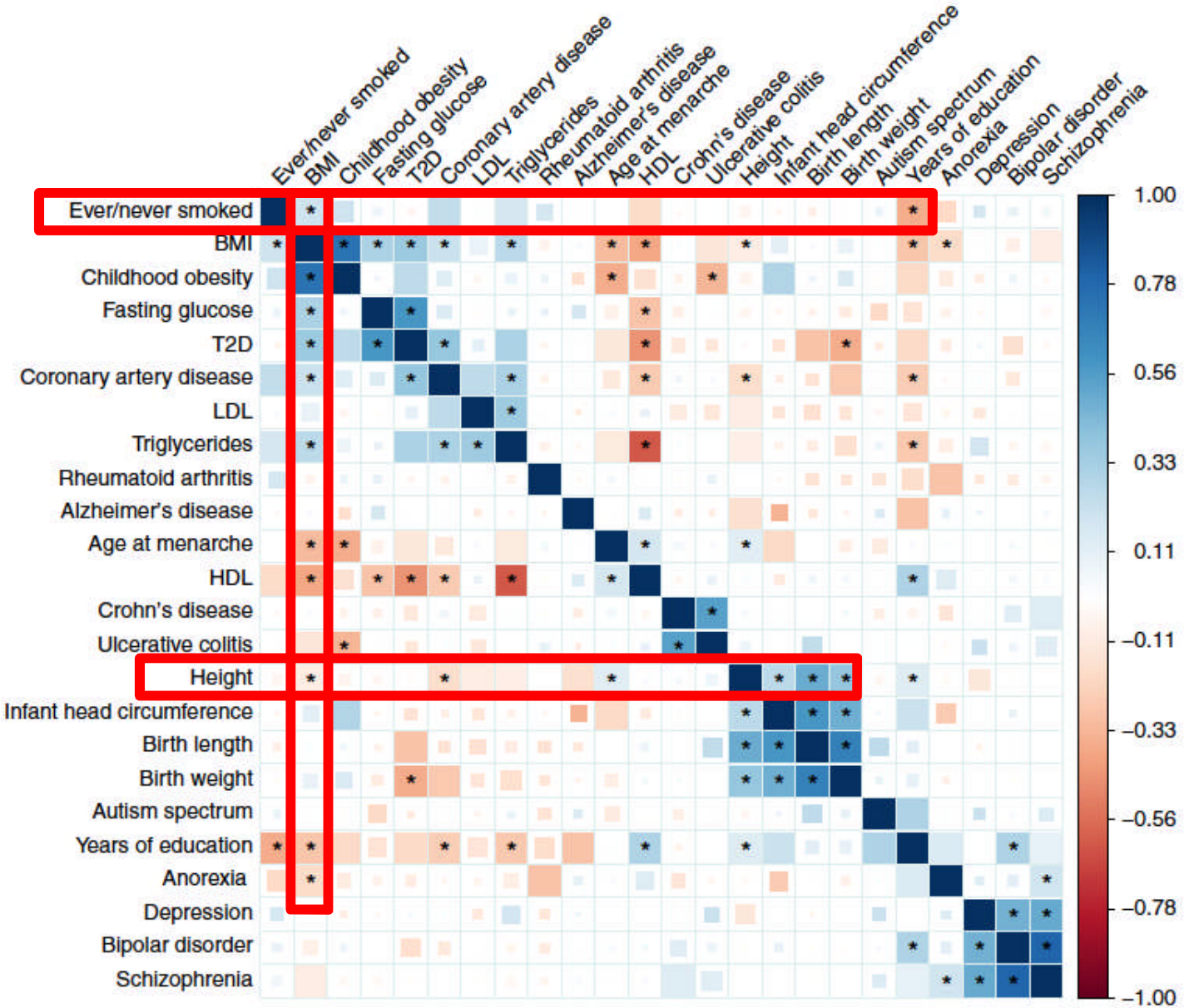
- Builds on Quantitative Genetics research to tell us:
  - What are the specific genetic factors?
  - How big is their effect?
  - How do they influence human trait variation or liability to disease?

Why do we estimate heritability?

- Quantitative Genetic studies tell us:
  - The importance of genetic and environmental influence
  - Provides a best case scenario for genetic prediction
  - Allows an estimate the success of gene identification efforts

- They do not tell us:
  - Number of genes
  - Where they are located
  - Mechanism of action
  - Magnitude of effect
  - Types of genes

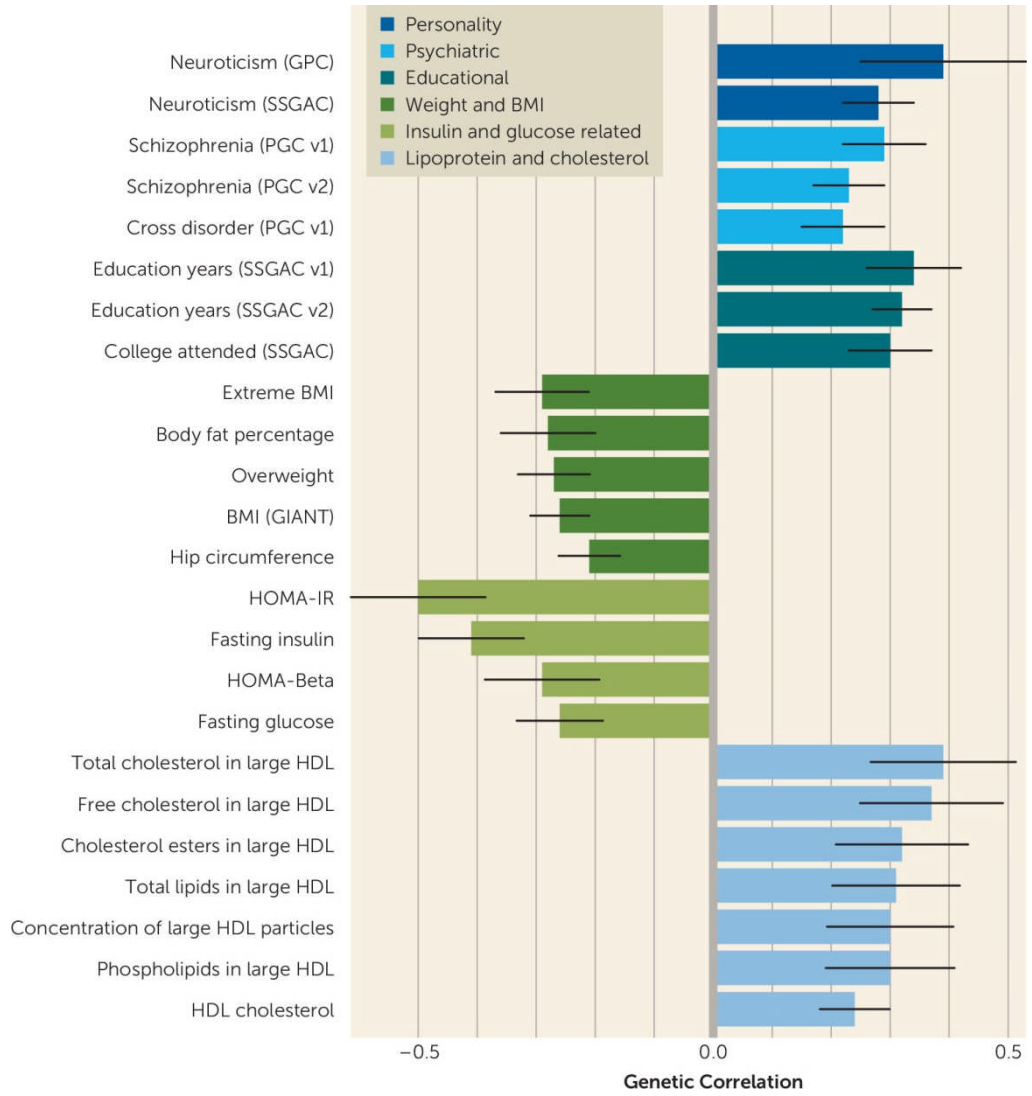
An atlas of genetic correlations across human diseases and traits



NATURE GENETICS
NOVEMBER 2015

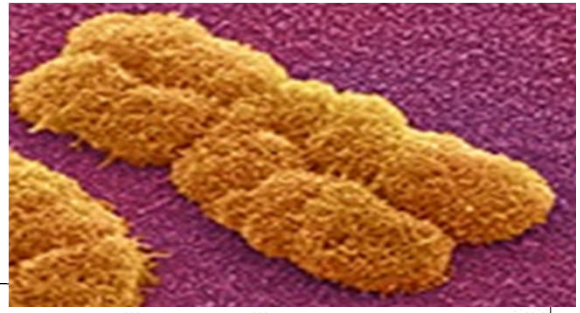
PMID: 26414676

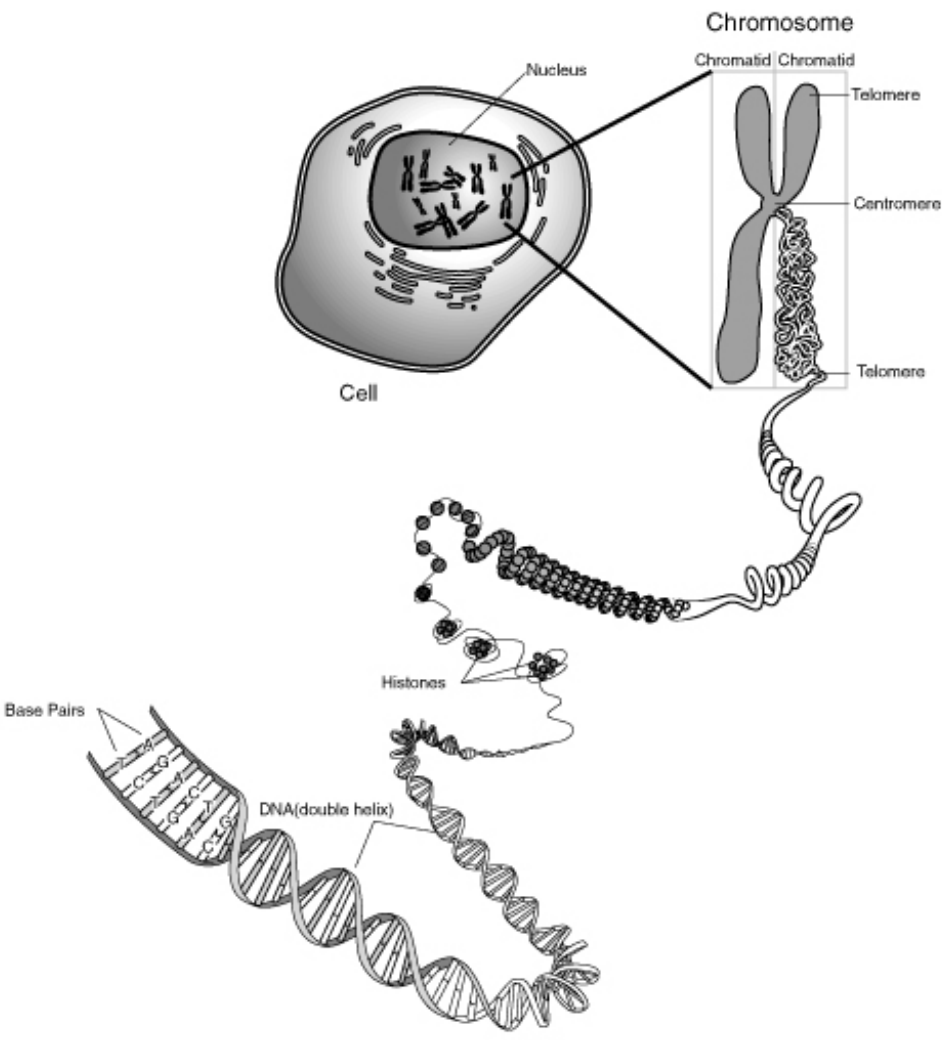
GWAS of Anorexia 2017

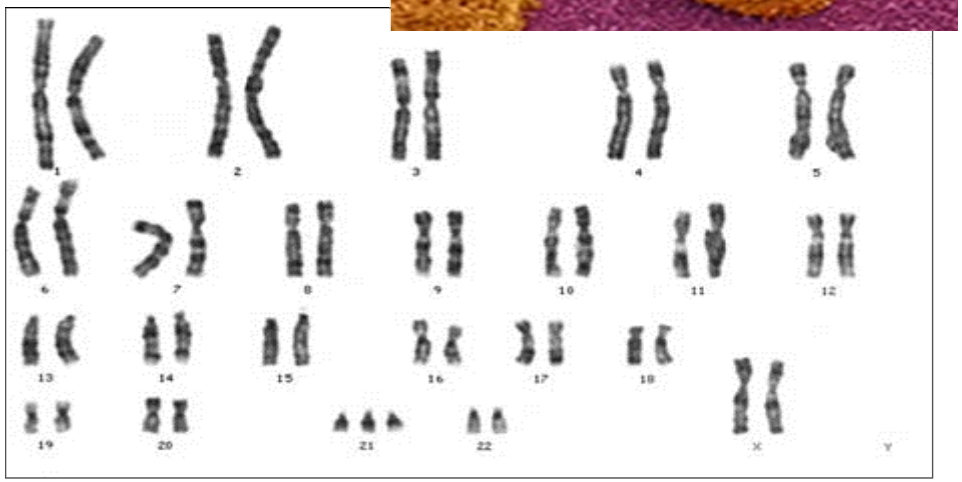


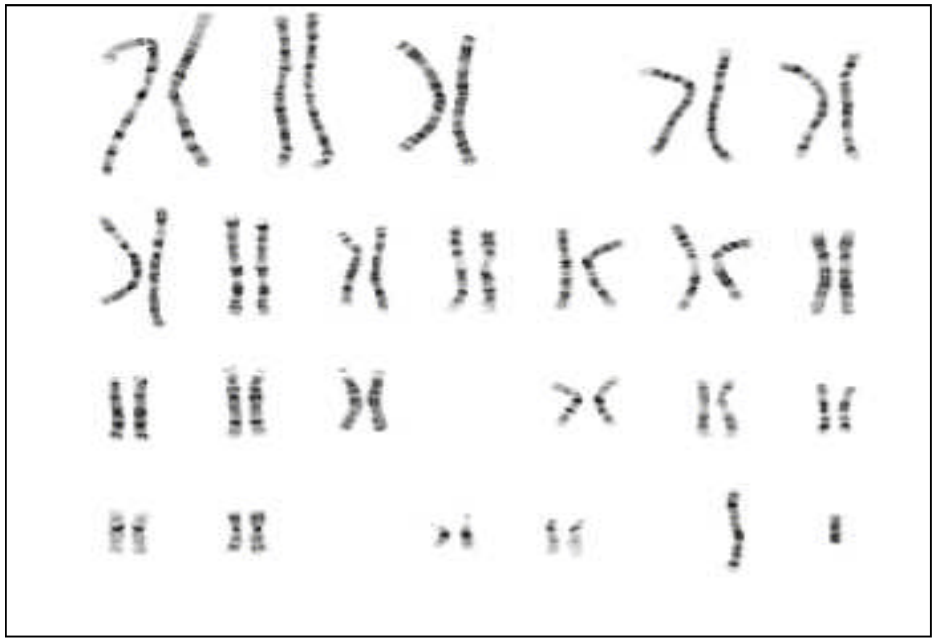
3,495 cases and 10,982 controls; Duncan et al 2017 Am J Psych

DNA Structure

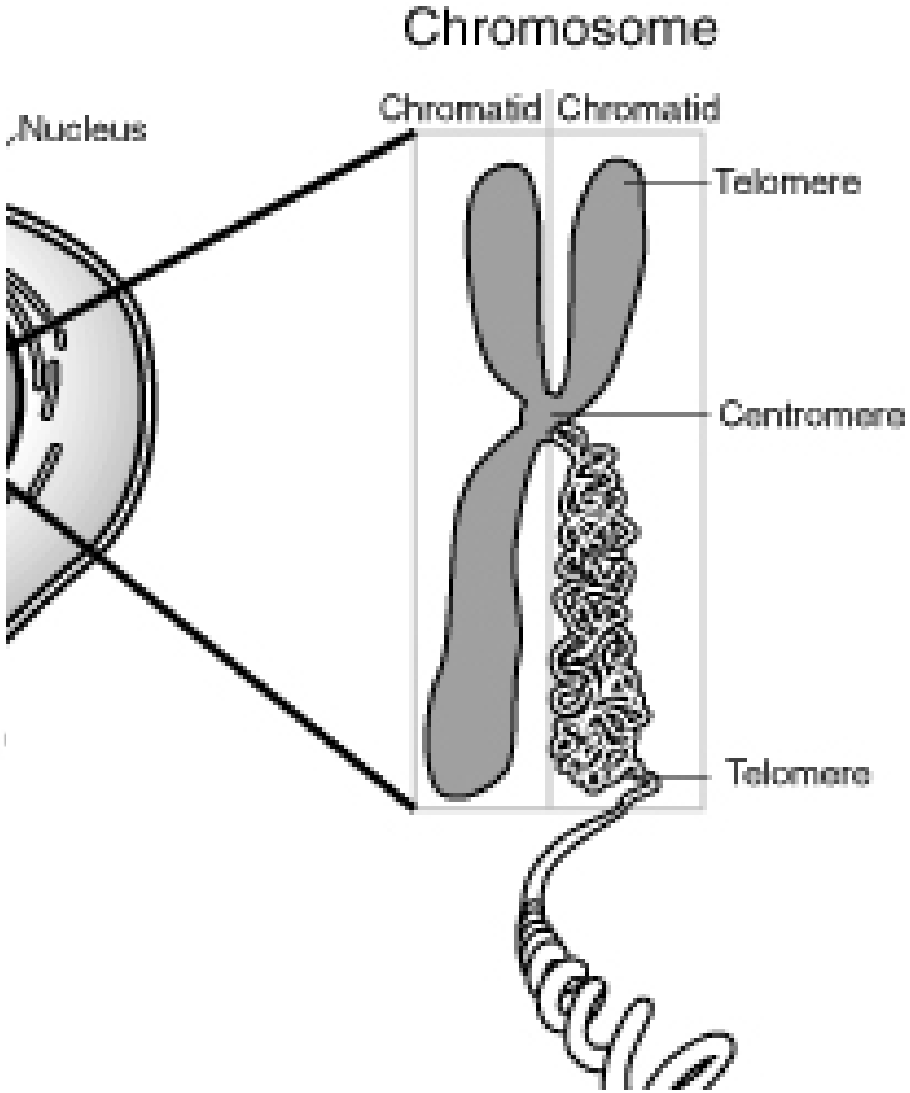




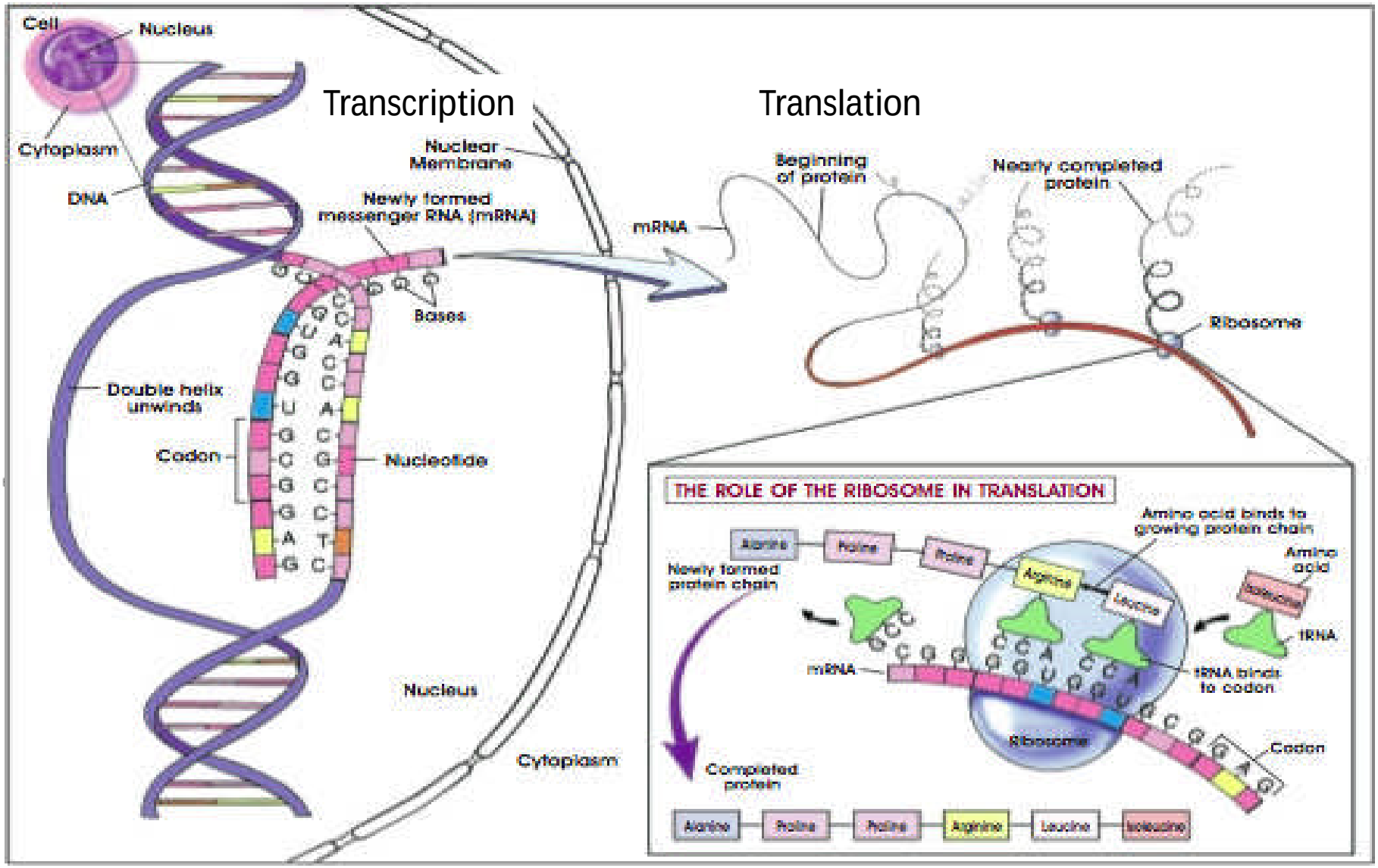




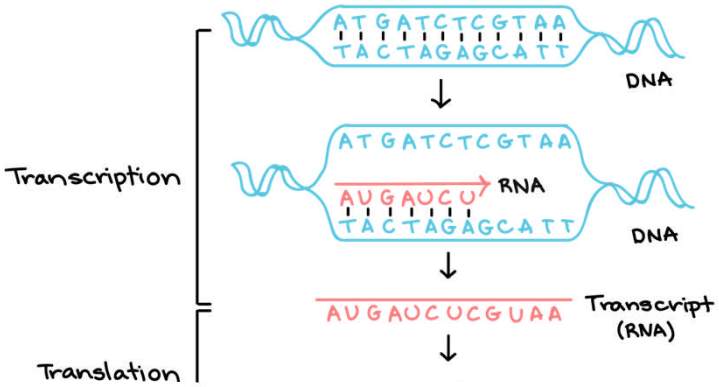
DNA Structure

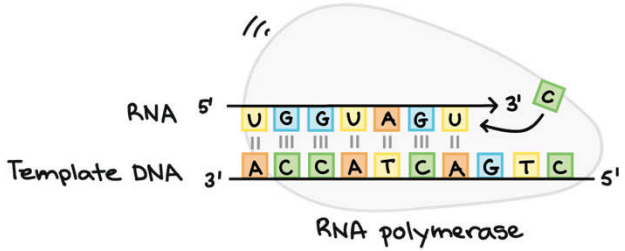


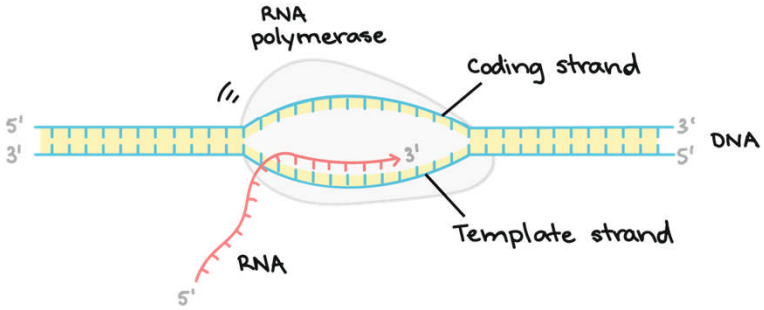
DNA to mRNA to Protein

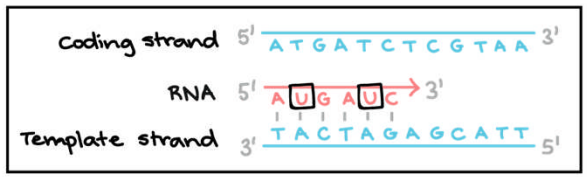


Transcription (DNA to RNA)

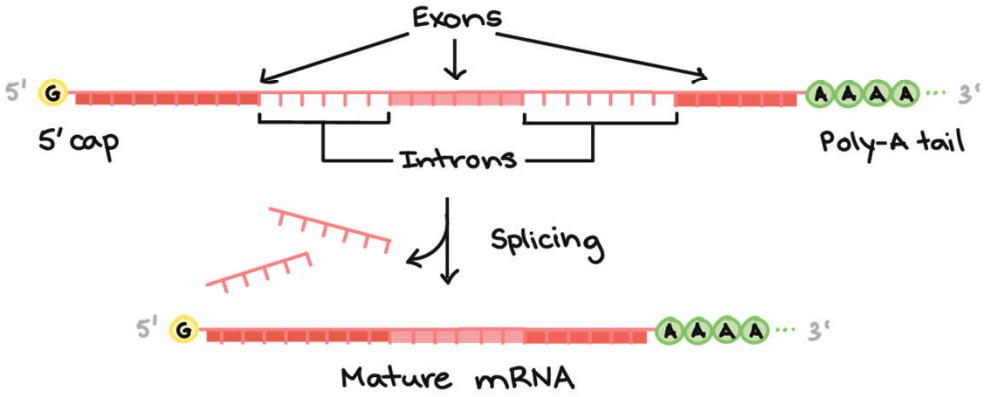






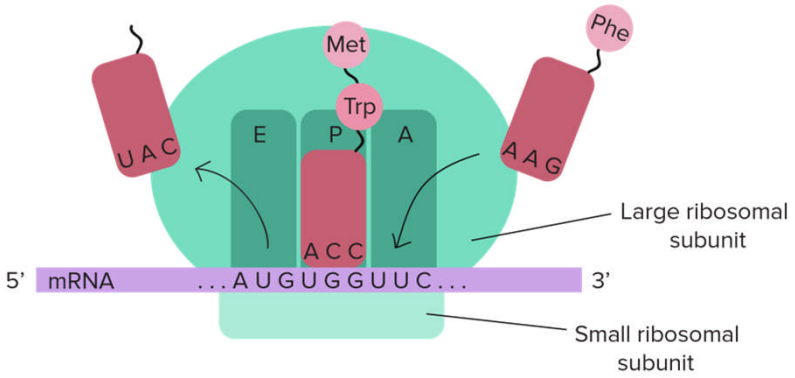


- The process of producing RNA from the DNA
- For each gene this is only from one strand of the DNA
- The pre-mRNA is "spliced" to form mature mRNA (removal of introns)



Images: https://ka-perseus-images.s3.amazonaws.com/20ce29364b2e7ff0cdea72acaa5b1dbd7287ab00.png, https://ka-perseus-images.s3.amazonaws.com/1e40590670cfdc967a5fe54d2e204df30e761232.png, https://ka-perseus-images.s3.amazonaws.com/1da89713b9aa8067742244d916749e72561bb3cc.png, https://ka-perseus-images.s3.amazonaws.com/63dbd2d4d2bbdba7861bd7d904113e4364adf71b.png

Translation



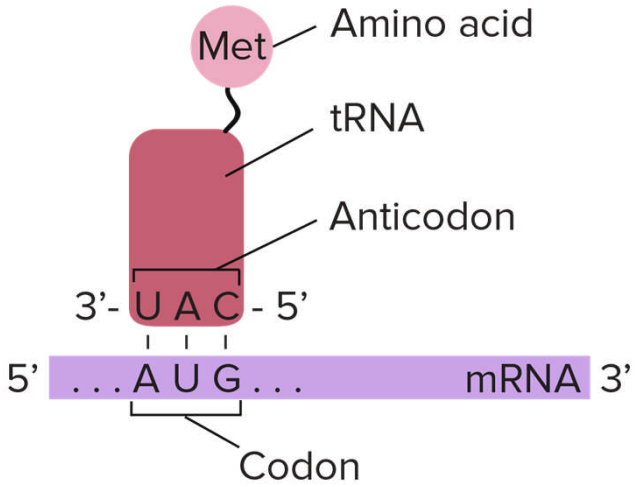


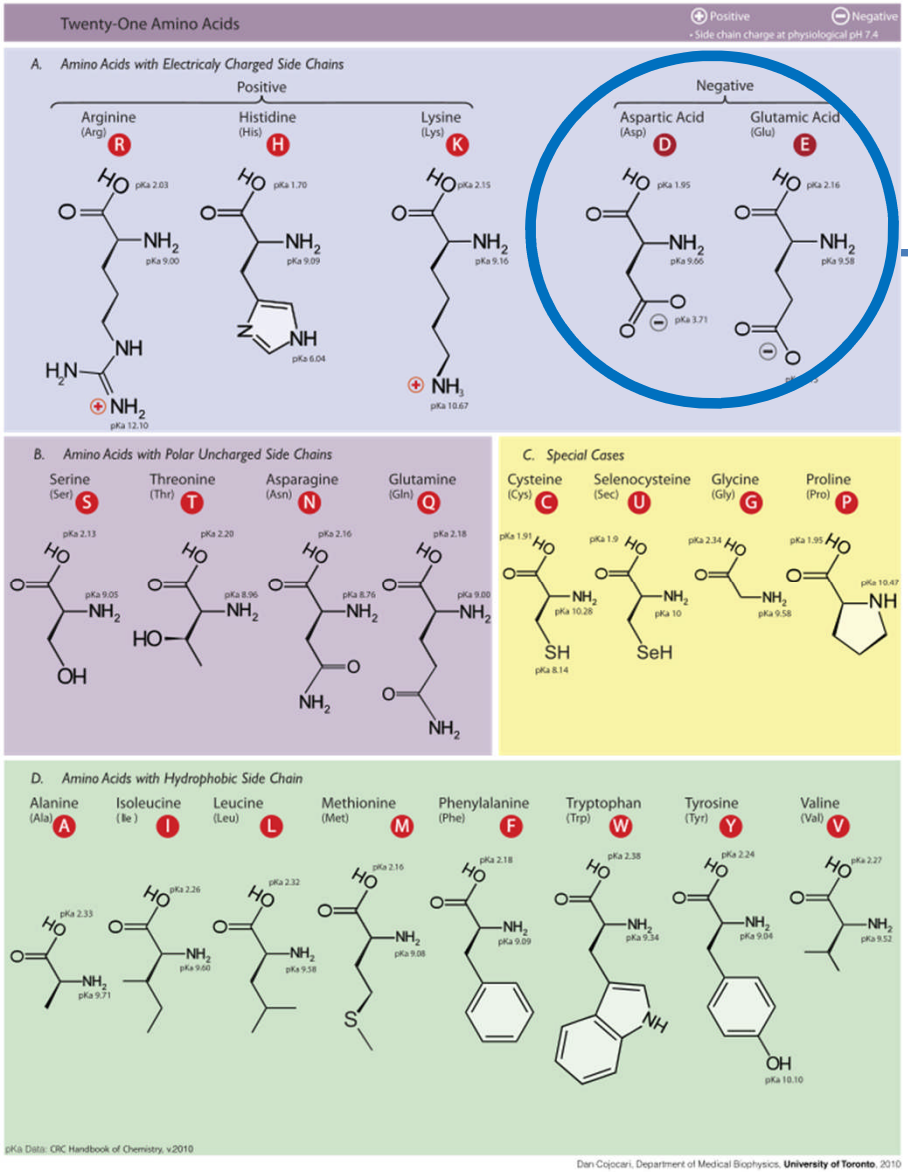
Image:
https://ka-perseus-images.s3.amazonaws.com/282fd6184d65eaf8e8284edccf3aa650ad11d774.png &
https://ka-perseus-images.s3.amazonaws.com/c5957e0217ce7123259c1918c2f8b337b08783a0.png

Second letter

First letter	U	C	A	G	Third letter
U	UUU Phe UUC Phe UUA Leu UUG Leu	UCU Ser UCC Ser UCA Ser UCG Ser	UAU Tyr UAC Tyr **UAA Stop** **UAG Stop**	UGU Cys UGC Cys **UGA Stop** UGG Trp	U C A G
C	CUU Leu CUC Leu CUA Leu CUG Leu	CCU Pro CCC Pro CCA Pro CCG Pro	CAU His CAC His CAA Gln CAG Gln	CGU Arg CGC Arg CGA Arg CGG Arg	U C A G
A	AUU Ile AUC Ile AUA Ile **AUG Met**	ACU Thr ACC Thr ACA Thr ACG Thr	AAU Asn AAC Asn AAA Lys AAG Lys	AGU Ser AGC Ser AGA Arg AGG Arg	U C A G
G	GUU Val GUC Val GUA Val GUG Val	GCU Ala GCC Ala GCA Ala GCG Ala	GAU Asp GAC Asp GAA Glu GAG Glu	GGU Gly GGC Gly GGA Gly GGG Gly	U C A G

Several codons code for the same amino acid
“Third Base Wobble”
This means that some DNA variants in coding regions do not lead to an amino acid change

Translation

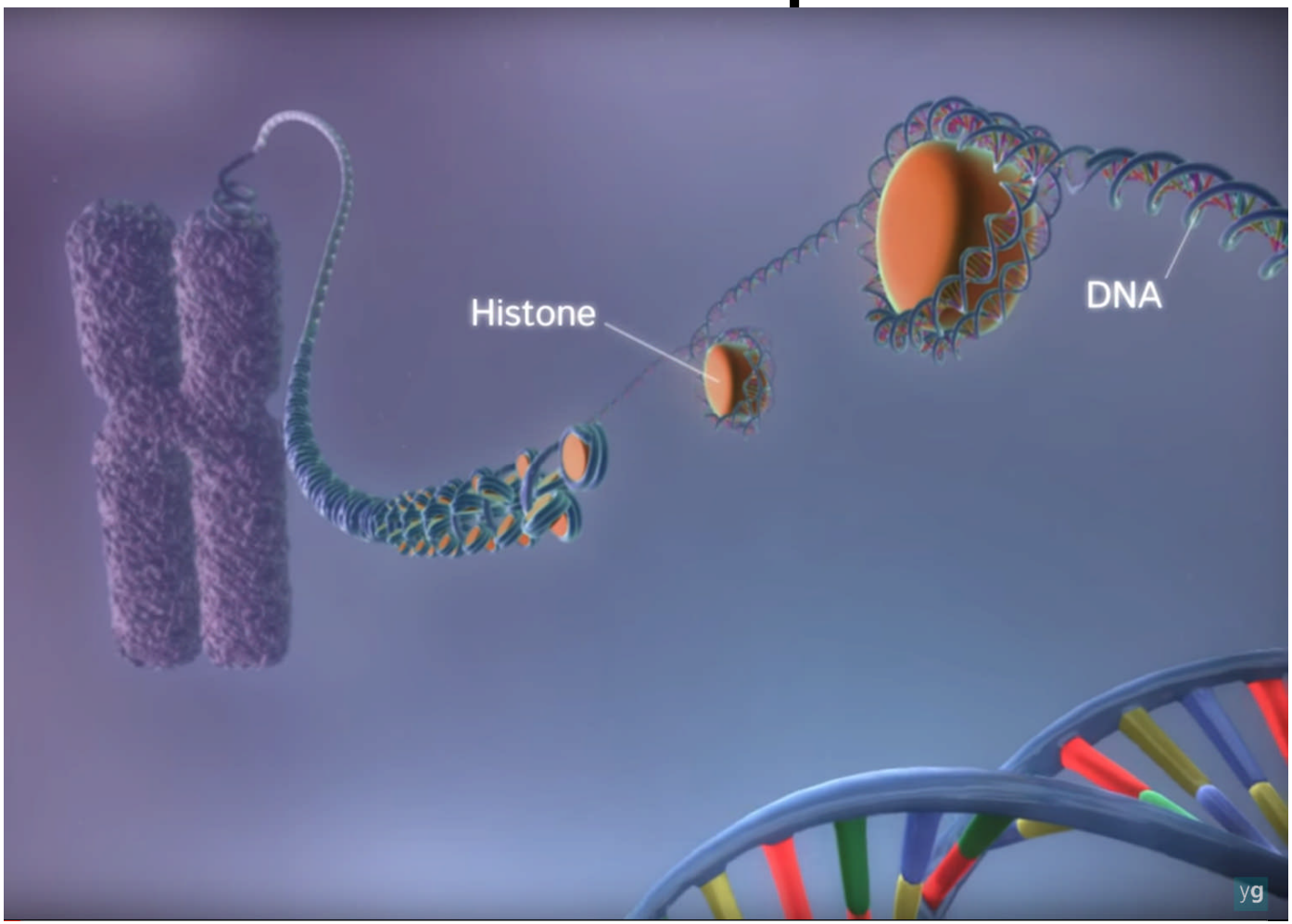


Second letter

First letter	U	C	A	G	Third letter
U	UUU Phe UUC UUA Leu UUG	UCU UCC UCA UCG Ser	UAU Tyr UAC **UAA Stop** **UAG Stop**	UGU Cys UGC **UGA Stop** UGG Trp	U C A G
C	CUU CUC CUA CUG Leu	CCU CCC CCA CCG Pro	CAU His CAC CAA Gln CAG	CGU CGC CGA CGG Arg	U C A G
A	AUU AUC AUA Ile AUG Met	ACU ACC ACA ACG Thr	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
G	GUU GUC GUA GUG Val	GCU GCC GCA GCG Ala	GAU Asp GAC GAA Glu GAG	GGU GGC GGA GGG Gly	U C A G

Amino acids can be grouped into chemically or structurally similar groups.
Some DNA variation that leads to an amino acid change has little or no impact on protein structure/function

From DNA to protein



http://www.yourgenome.org/video/from-dna-to-protein
https://www.youtube.com/watch?v=gG7uCskUOrA
Slightly more “out there”: https://www.youtube.com/watch?v=-ygpqVr7_xs&t=3s

Why are gene based studies important? Genes to Behaviour

- Fundamental insights into biology of traits and disease
- Diagnosis

– Gene-based (rather than symptom-based) diagnoses.

– Disease heterogeneity.

Same aetiology, different phenotype

– Genetic heterogeneity.

Different aetiology, same phenotype

- Prevention (more likely, prediction)

– DNA as early warning system for behavioural and environmental (not genetic engineering)

- Treatment

– Allow new directions for therapy

– Tailored to individuals

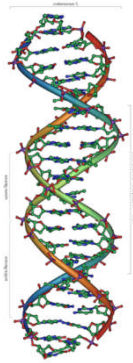
(e.g., pharmacogenetics)

– Relevant environments

(e.g., Tailoring Of Lifestyle choices)

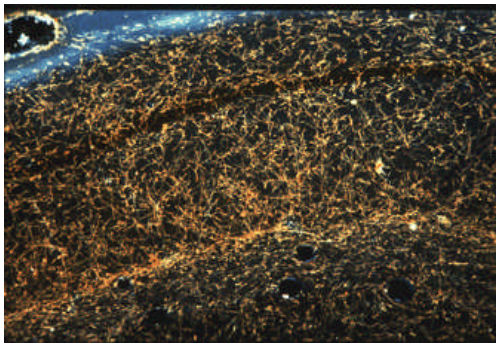
Why are gene based studies important?
Genes to Behaviour

Genes

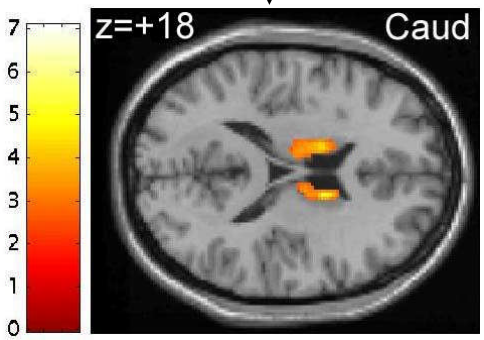


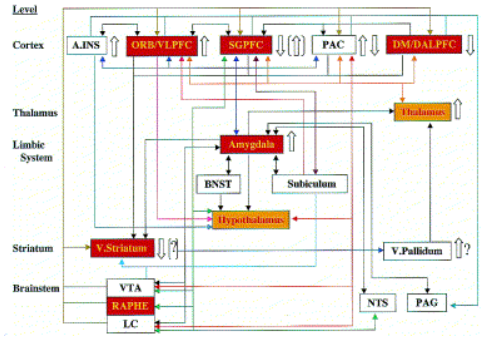
Proteins

Cells



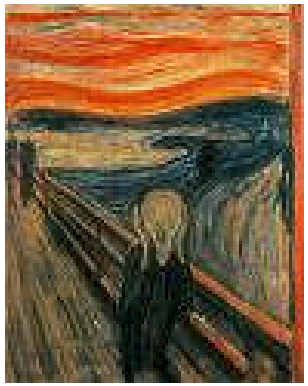
Brain regions





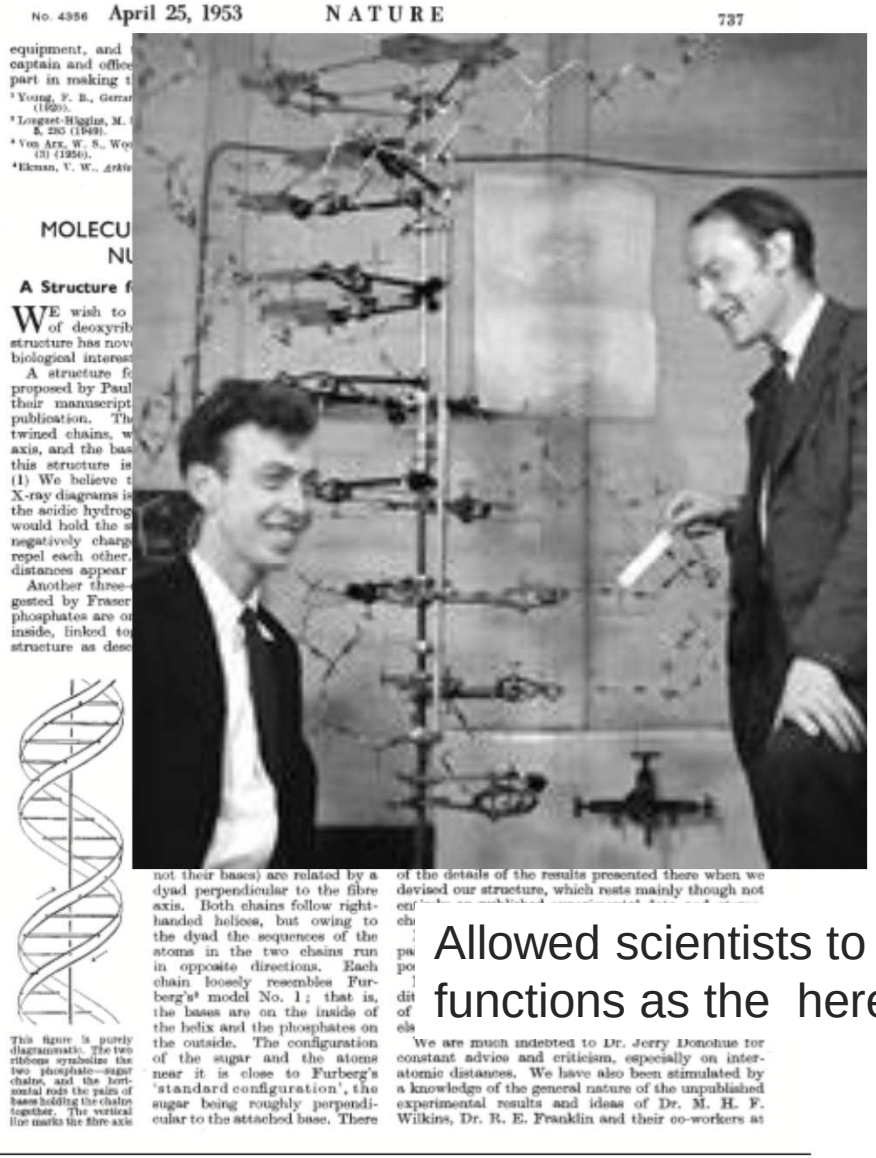
Brain systems

Cognition/ emotion



Slide Jon Roiser UCL

Discovery of the Structure of DNA



No. 4356 April 25, 1953 NATURE 737

equipment, and
captain and office
part in making t

[1] Young, F. B., Germ
(1950).
[2] Longuet-Higgins, M.
5, 285 (1949).
[3] Von Arx, W. S., Wo
(3) (1956).
[4] Ekman, V. W., Arkiv

MOLECU
NU

A Structure f

WE wish to
of deoxyrib
structure has nov
biological interes
A structure f
proposed by Paul
their manuscript
publication. The
twined chains, w
axis, and the bas
this structure is
(1) We believe t
X-ray diagrams i
the acidic hydrog
would hold the s
negatively charg
repel each other.
distances appear
Another three-
gested by Fraser
phosphates are o
inside, linked to
structure as desc

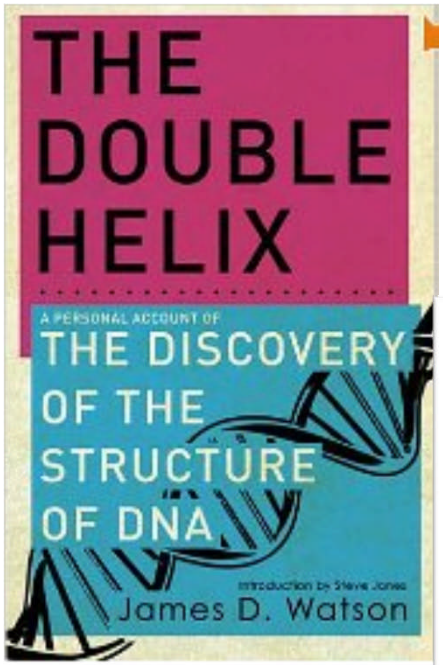
This figure is purely diagrammatic. The two ribbons symbolize the two phosphate–sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis.

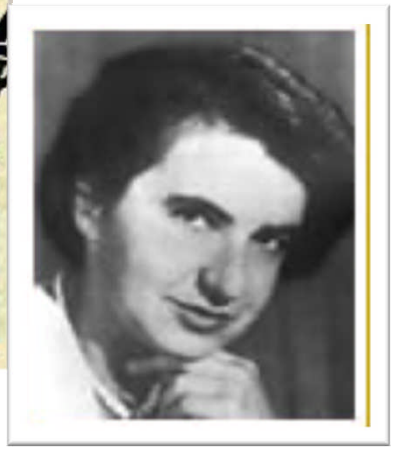
not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There

of the details of the results presented there when we devised our structure, which rests mainly though not

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

Watson & Crick: 1953

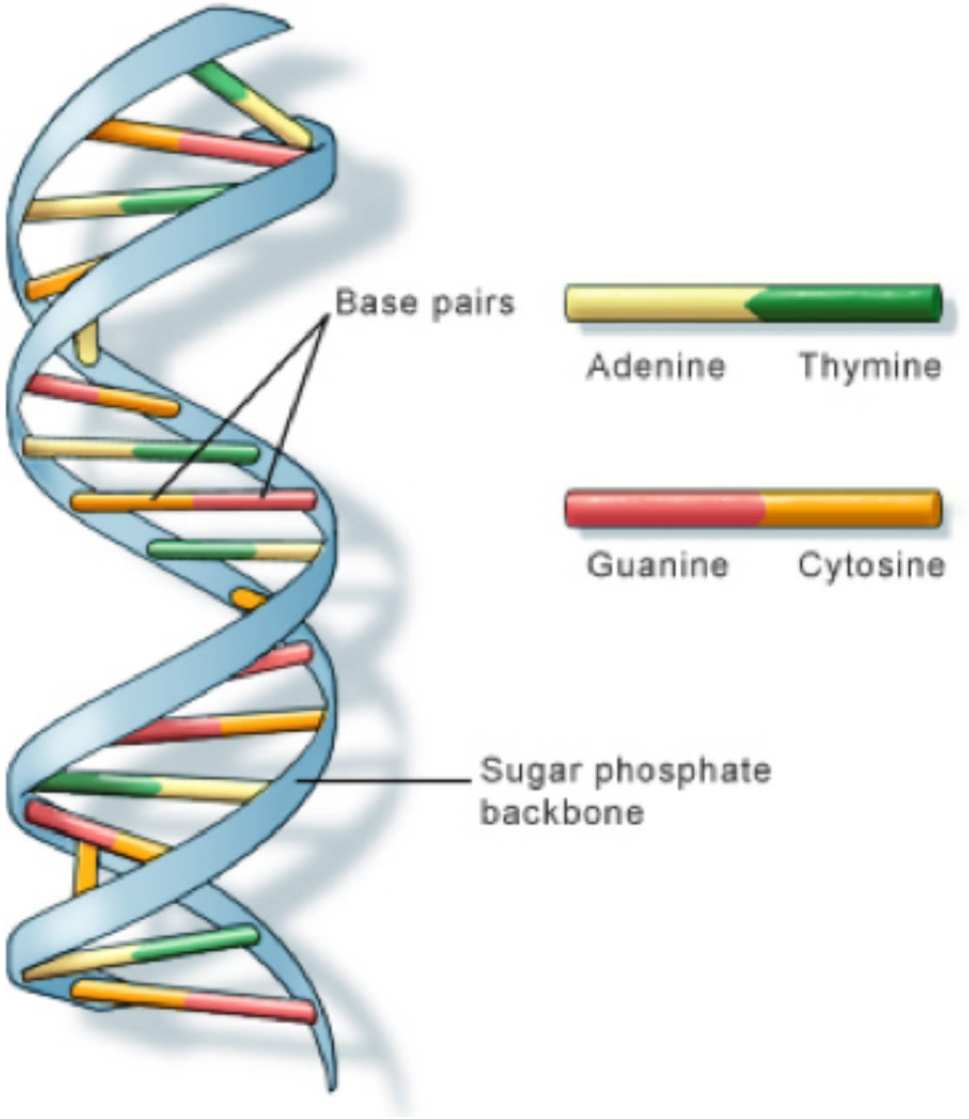




Allowed scientists to ask how DNA functions as the hereditary material

Rosalind Franklin:
X-ray images of DNA

DNA Structure



DNA: Deoxyribonucleic Acid
Structure: double helix

DNA molecule is made up of *sugar* residues *phosphate* groups and *bases*

Attached to carbon atom 1' of each sugar is a nitrogenous base:
Adenine (A)
Cytosine (C)
Guanine (G)
Thymine (T)

Watson-Crick base-pairing rules:
A:T, C:G (hydrogen bonds)

The Human Genome Project

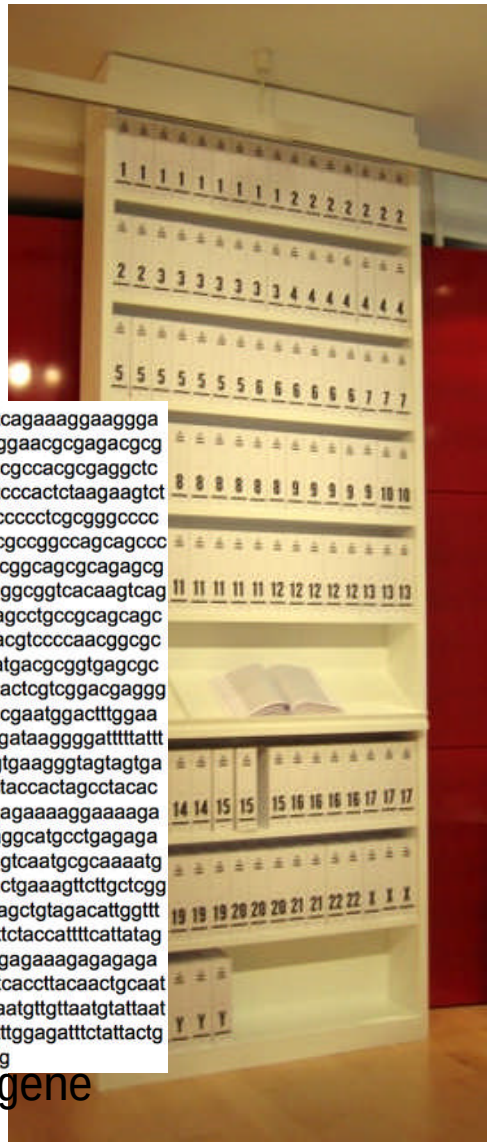
Size of the Human Genome

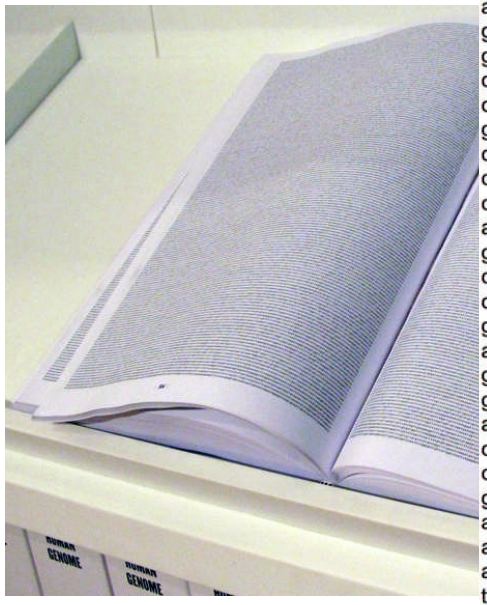
The Number of Genes

Variation between individuals

The Human Genome

- Published in 2003
- 3.1×10^9 bases (3 billion)
- 22,000 protein coding genes (~2%)
- $3 billion USD, 15 years





agcactctctcacttctggccagggaacgtggaaggcgcaccgacagggatccggccagggagggcgagtgaaagaaggaaatcagaaaggaaggga
gttaacaaaataataaaaacagcctgagccacggctggagagaccgagacccggcgcaagagagcgcagccttagtaggagaggaacgcgagacgcg
gcagagcgcgttcagcactgacttttgctgctgcttctgctttttttttcttagaaacaagaaggcgccagcggcagcctcacacgcgagcgccacgcgaggctc
ccgaagccaacccgcgaagggaggaggggagggaggaggaggcggcgtgcagggaggagaaaaagcattttcacttttttgctcccactctaagaagtct
cccggggattttgtatatatttttttaacttccgtcagggctcccgcttcatatttccttttctttccctctctgttcctgcacccaagttctctctgtgtccccctcgcgggcccc
gcacctcgcgtcccggatcgctctgattccgcgactccttggccgccgctgcgcatggaaagctctgccaagatggagagcggcggcgccggccagcagccc
cagccgcagcccccagcagcccttcctgccgcccgcagcctgtttctttgccacggccgcagccgcggcggccgcagccgccgcagcggcagcgcagagcg
cgcagcagcagcagcagcagcagcagcagcagcagcaggcgccgcagctgagaccggcggccgacggccagccctcaggggggcggtcacaagtcag
cgcccaagcaagtcaagcgacagcgctcgtcttcgcccgaactgatgcgctgcaaacgccggctcaacttcagcggctttggctacagcctgccgcagcagc
agccggccgccgtggcgcgccgcaacgagcgcgagcgcaaccgcgtcaagttggtcaacctgggctttgccacccttcgggagcacgtccccaacggcgc
ggccaacaagaagatgagtaaggtggagacactgcgctcggcggtcgagtacatccgcgcgctgcagcagctgctggacgagcatgacgcggtgagcgc
cgccttccaggcaggcgtcctgtcgcccaccatctcccccaactactccaacgacttgaactccatggccggctcgccggtctcatcctactcgtcggacgaggg
ctcttacgacccgctcagccccgaggagcaggagcttctcgacttcaccaactggttctgaggggctcggcctggtcaggccctggtgcgaatggacttggaa
gcaggtaggttgcattttggggtgggcaggggggtattcttgccttcgtcctccctctgagtgtctgtggaagtggggatgtctccaaggagataaggggatttttatttt
aaagaatttgtgaaagttggtcgatttcaagtcctagtttgttagtttcagcactggcctctgaaaatggccttgcccaggtctccaaggagtgaagggtagtagtga
ggtgcagagatactggtgaaccgaatactgggacatgttaaaagagatgtctacctgacagactctttccccagacctccatctccctctaccactagcctacac
gttcaaattaacctctcctgttcttttccttatgttatagggtgatcgcacaacctgcatctttagtgctttcttgtcagtggcgttgggaggggggagaaaaggaaaaga
aaaaaaaaagaagaagaagaagaaaagagaagaagaaaaaaacgaaaacagtcaaccaaccccatcgccaactaagcgaggcatgcctgagaga
catggctttcagaaaacgggaagcgctcagaacagtatctttgcactccaatcattcacggagatatgaagagcaactgggacctgagtcaatgcgcaaaatg
cagcttgtgtgcaaaagcagtgggctcctggcagaagggagcagcacacgcgttatagtaactcccatcacctctaacacgcacagctgaaagttcttgctcgg
gtcccttcacctcctcgcccttttcttaaagtgcagttcttagccctctagaaacgagttggtgtctttcgtctcagtagcccccaccccaataagctgtagacattggttt
acagtgaaactatgctattctcagccctttgaaactctgcttctcctccagggcccgattcccaaaccccatggcttccctcacactgtcttttctaccattttcattatag
aatgcttccaatctttgtgaatttttattataaaaaatctatttgtatctatcctaaccagttcggggatatattaagatatttttgtacataagagagaaagagagaga
aaaatttatagaagttttgtacaaatggtttaaaatgtgtatatcttgatactttaacatgtaatgctattacctctgcatattttagatgtgtagttcaccttacaactgcaat
tttccctatgtggttttgtaaagaactctcctcataggtgagatcaagaggccaccagttgtacttcagcaccaatgtgtcttactttatagaaatgttgttaatgtattaat
gatgttattaaatactgttcaagaagaacaaagtttatgcagctactgtccaaactcaaagtggcagccagttggttttgataggttgccttttggagatttctattactg
ccttttttttttcttactgttttattacaaacttacaaaaatatgtataaccctgttttatacaaactagtttcgtaataaaacttttttcctttttttaaaatg

ASCL1 (Achaete-scute complex-like 1) gene

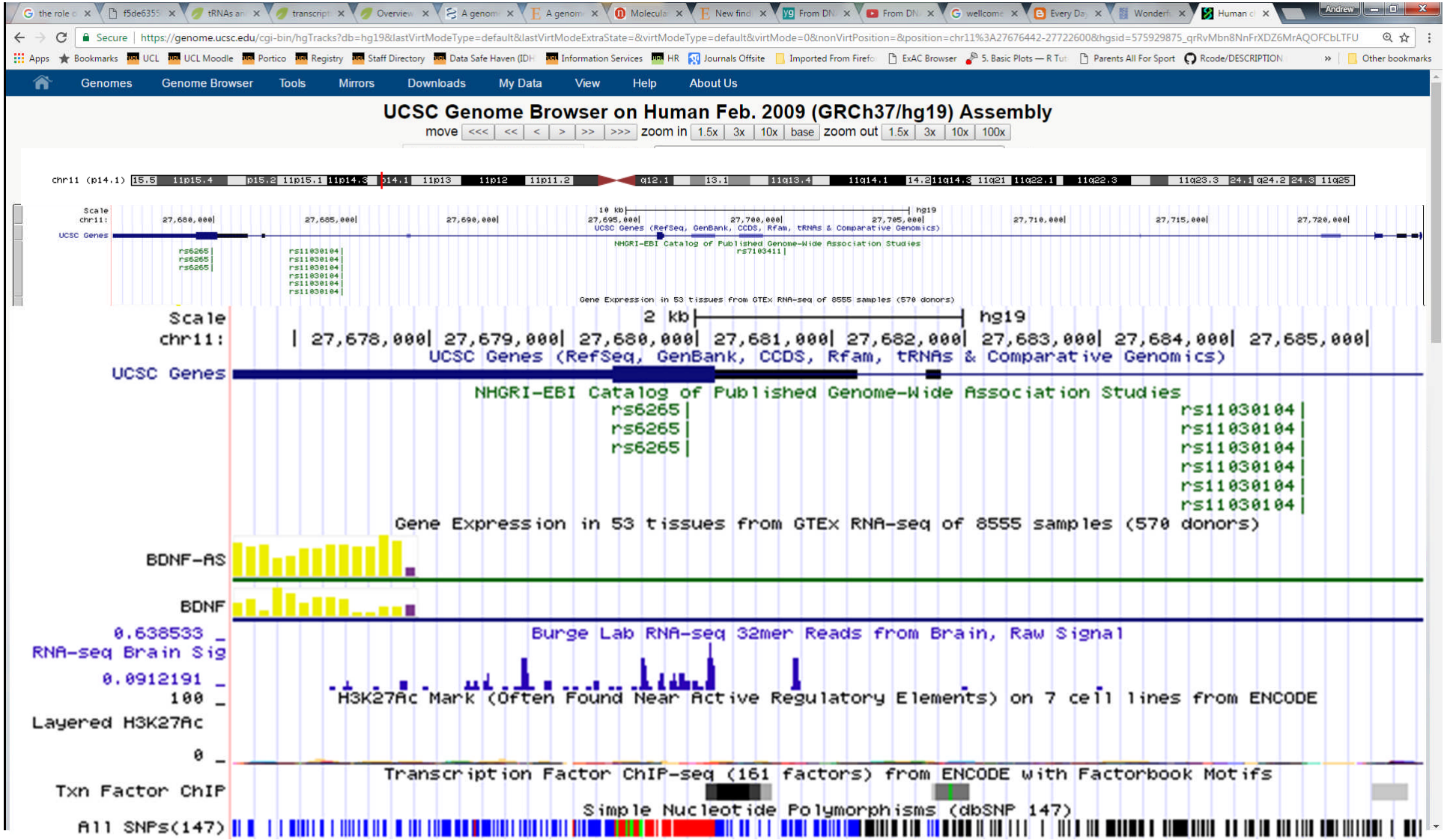
https://www.youtube.com/watch?v=yqCJLRsBsFw

Announcement of a working draft June 2000

Navigating the Genome

- Genome Browsers
  - Gene Annotations
  - Phenotype Associations
  - Gene Expression
  - Etc…

  - http://genome.ucsc.edu/
  - http://www.ensembl.org/index.html



http://genome.ucsc.edu/

We vary genetically

- No two people are genetically identical, except for MZ twins (with caveats)
- Humans are ~99.9% genetically identical
- We mostly know where the ~0.1% (3 million bases) of DNA variation resides
- Variation makes us unique

https://www.nigms.nih.gov/Education/pages/Factsheet_studyinggenes.aspx

Classification of Genetic Variation

- ‘Mutation’ Change in DNA sequence from wild type. Often used as short-hand for pathogenic variant
- ‘Polymorphism, or DNA variant’ A variation of the DNA sequence (at a specific locus) that is present in more than 1% of the population (*minor allele frequency*; MAF >1%)
- ‘Non-polymorphic’ The DNA base does not vary in the population

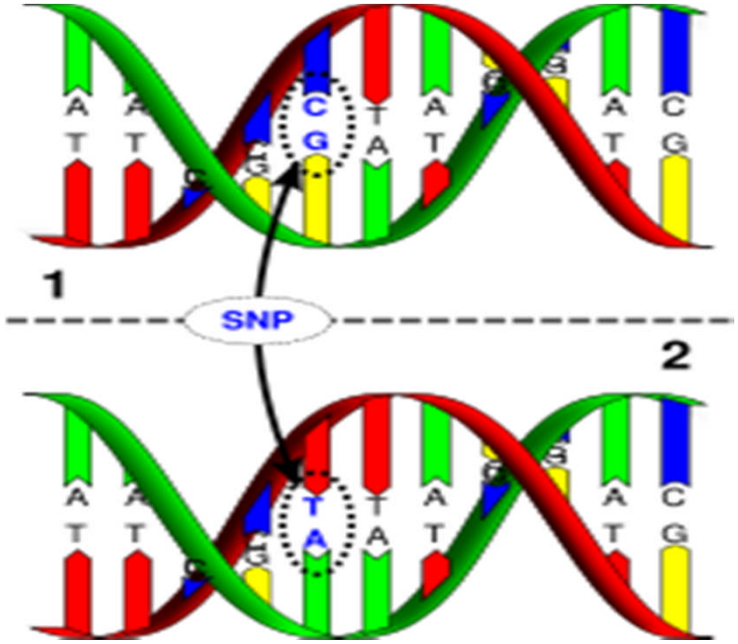
Glossary of Terms

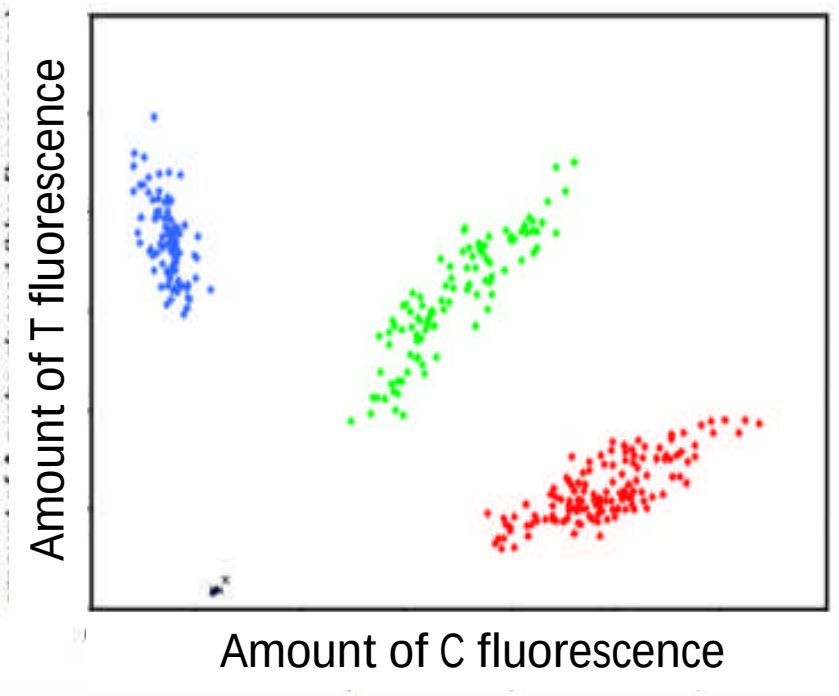
- Genes:
  - Units of inheritance that are passed down to the next generation
- Locus:
  - Specific position in the genome
- Alleles:
  - The genes found at the same locus on different homologous chromosomes are alleles.
  - *Allele frequency* refers to the number of alleles in the population
- Genotype:
  - Combination of the two alleles at a specific locus.
  - Alleles can be the same (*homozygous*) or different (*heterozygous*)

Many types of genetic variation

- Small scale:
  - Single base change (SNPs)
- Medium scale:
  - Copy number variation (CNVs), indels
- Large scale:
  - Chromosomal abnormalities
    - Change in the number of chromosomes (whole)
    - Change in the arrangement chromosomes (part)
- And everything in between…

SNPs: Alleles and Genotypes





Genotypes	CC	CT	TT
Controls	500	400	100
Cases	400	500	200

Genotypes to Alleles and MAFs

Genotypes	CC	CT	TT
Controls	500	400	100
Cases	400	500	200

Alleles	C		T	
Controls	1400	70%	600	30%
Cases	1300	59%	900	41%

TT individuals have two T alleles = 200 T alleles in the controls.
CT individuals have one T allele = 400 T alleles in the controls.
There are 600 T alleles in the controls.

The total number of allele in the controls is 2000 (each person has two alleles.
600/2000 = 30%.

Association Testing
Chi-square 54.303 $p=1.72x10^{-13}$
Odds ratio 1.615 (95% CI 1.421-1.836)

Genome Wide Association Studies

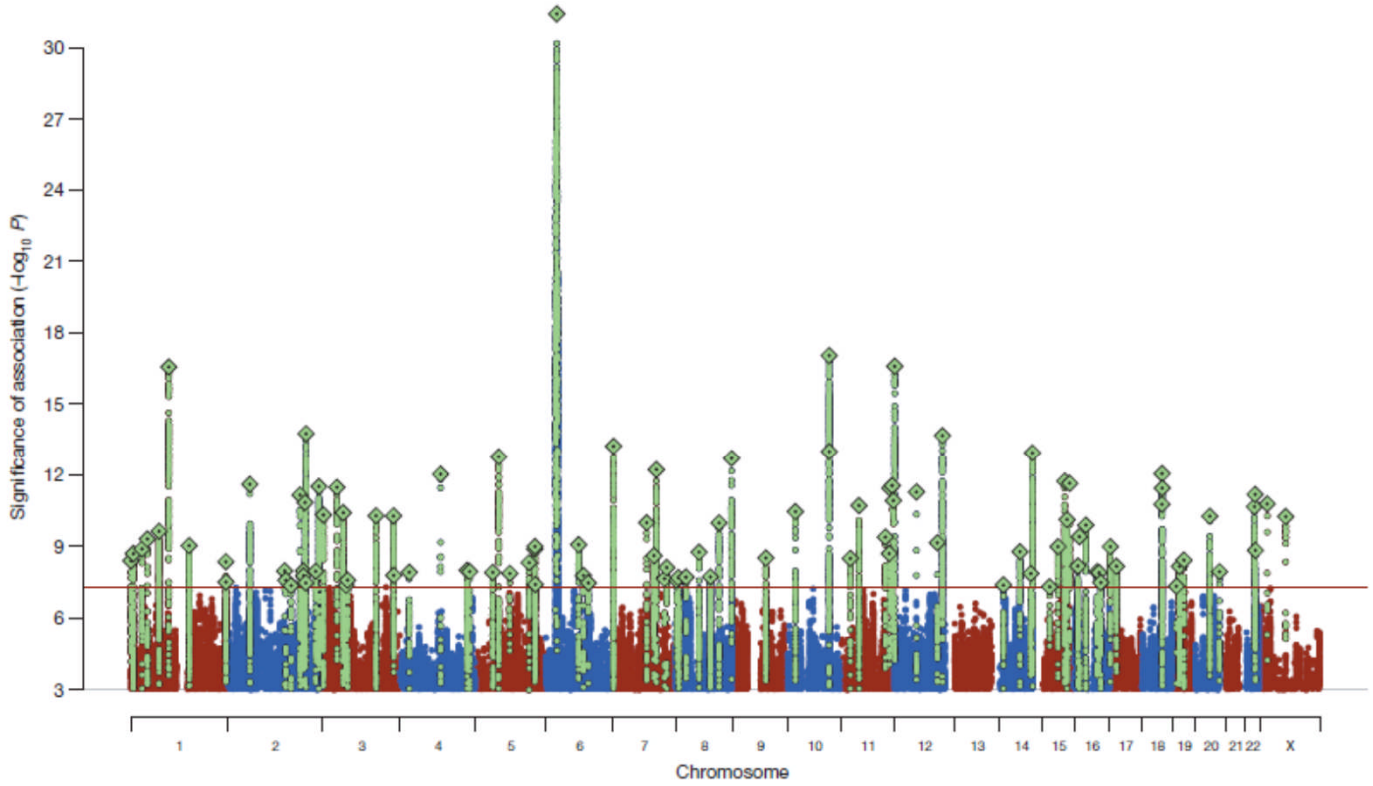
ARTICLE

doi:10.1038/nature13595

Biological insights from 108 schizophrenia-associated genetic loci

Schizophrenia Working Group of the Psychiatric Genomics Consortium*

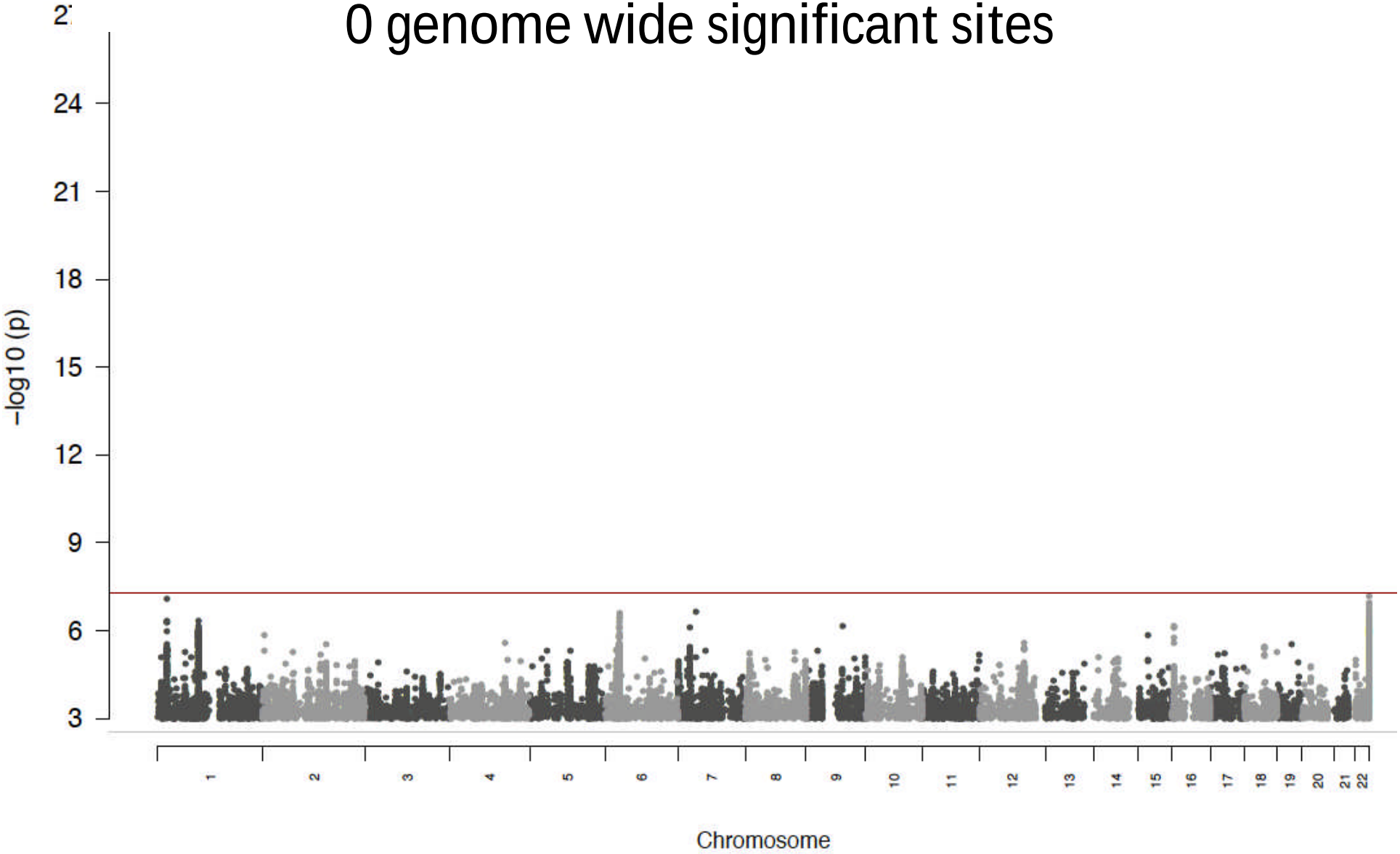
36,989 cases
113,075 controls
83 new loci
DRD2
several genes involved in glutamatergic neurotransmission



Schizophrenia GWAS 2009

2,601 cases, 3,345 controls

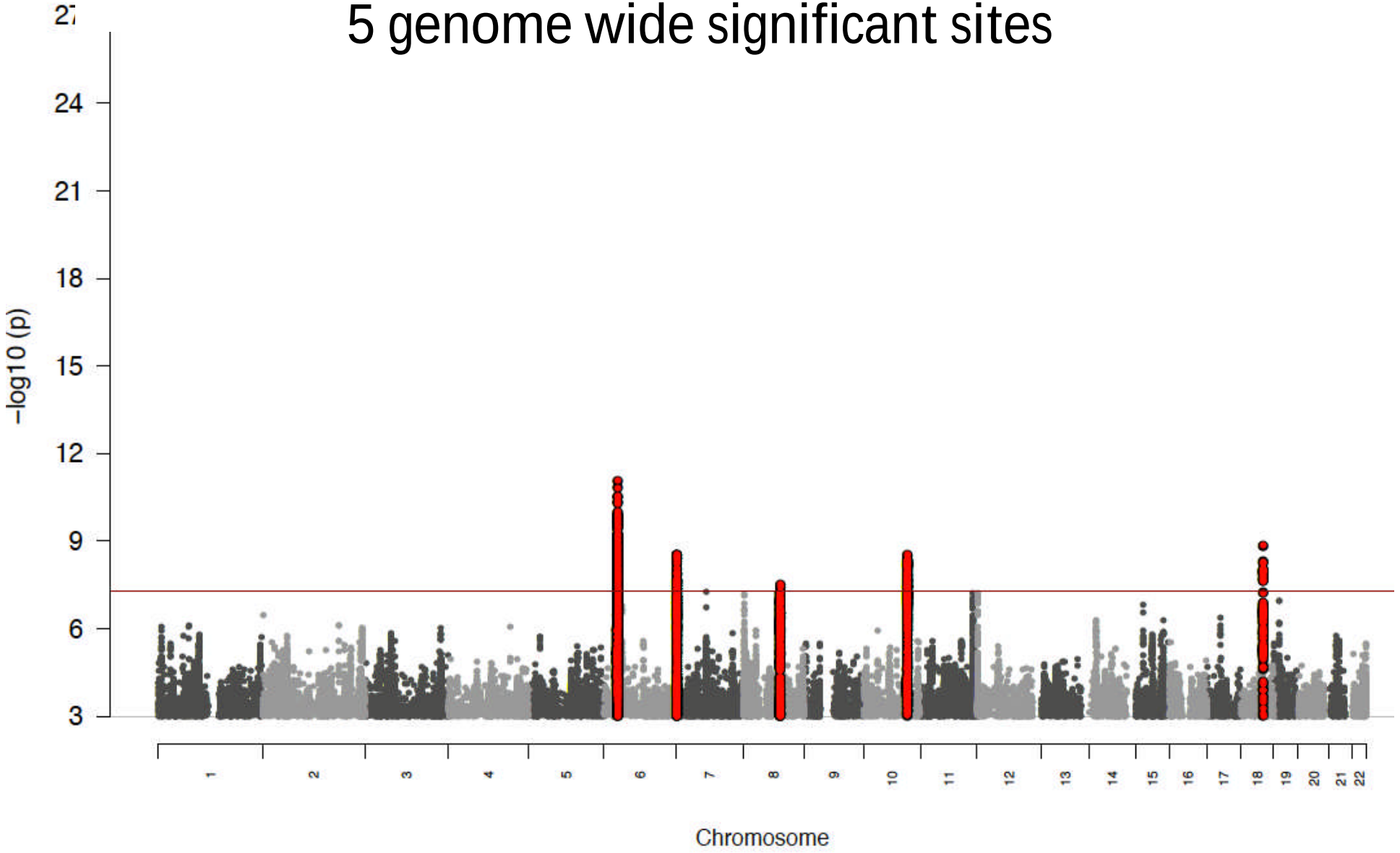
0 genome wide significant sites



Schizophrenia GWAS 2011

9,394 cases, 12,462 controls

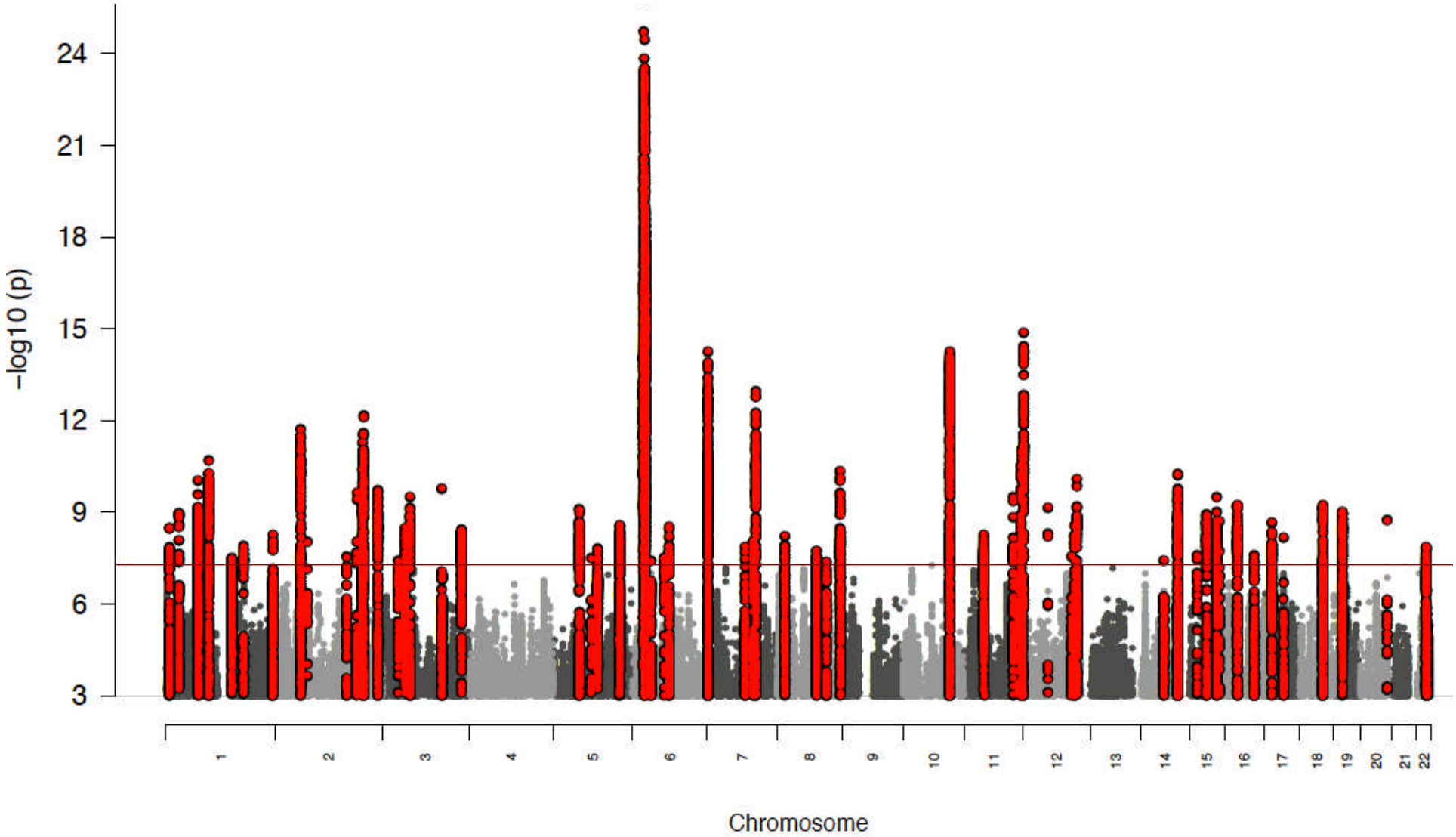
5 genome wide significant sites



Schizophrenia GWAS 2012

25,785 cases, 28,441 controls

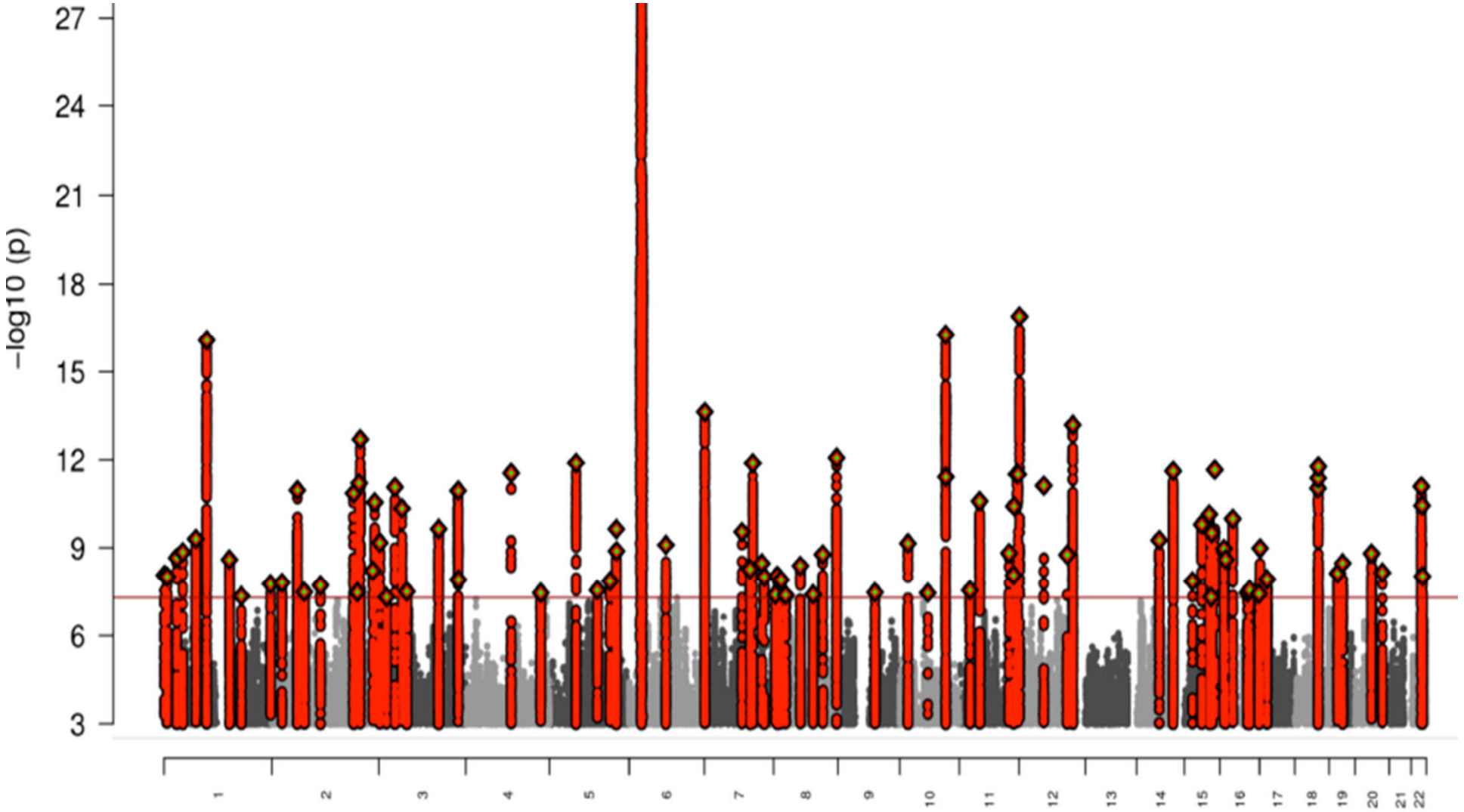
62 genome wide significant sites



Schizophrenia GWAS 2013

35,476 cases, 46,839 controls

97 genome wide significant sites



Genome Wide Association Studies

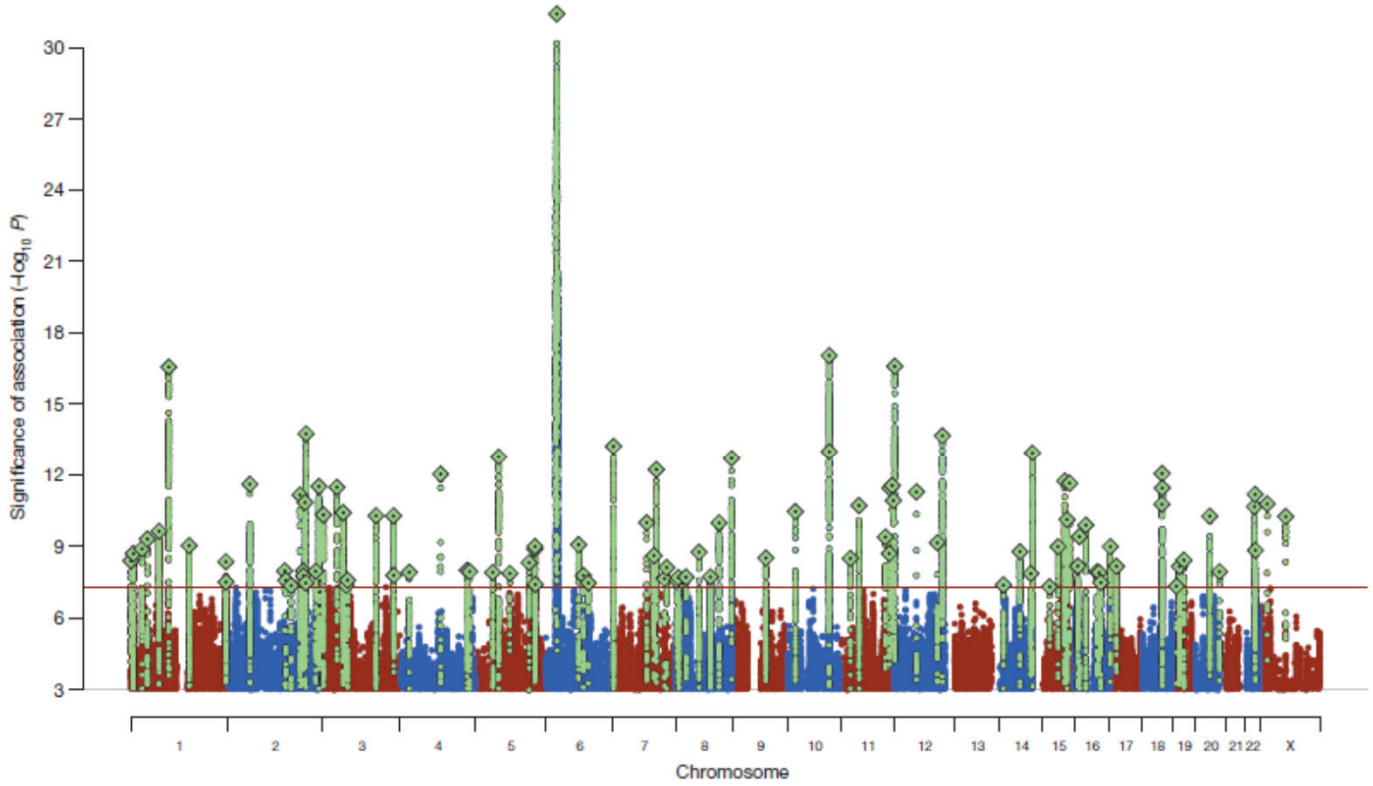
ARTICLE

doi:10.1038/nature13595

Biological insights from 108 schizophrenia-associated genetic loci

Schizophrenia Working Group of the Psychiatric Genomics Consortium*

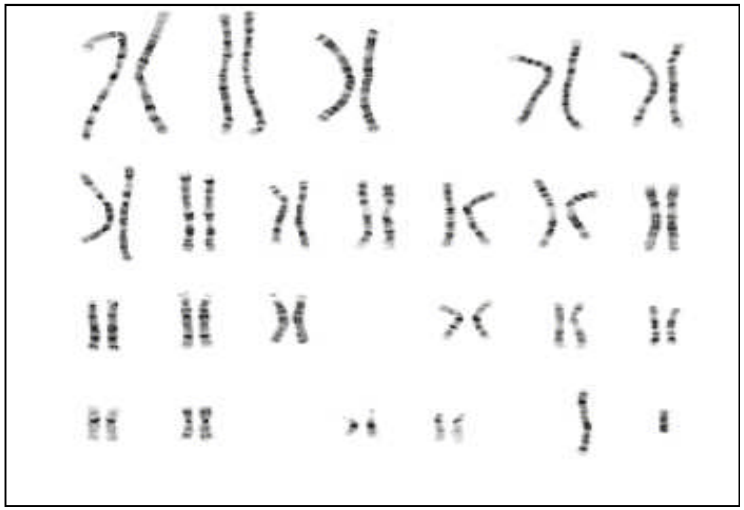
36,989 cases
113,075 controls
83 new loci
DRD2
several genes involved in glutamatergic neurotransmission

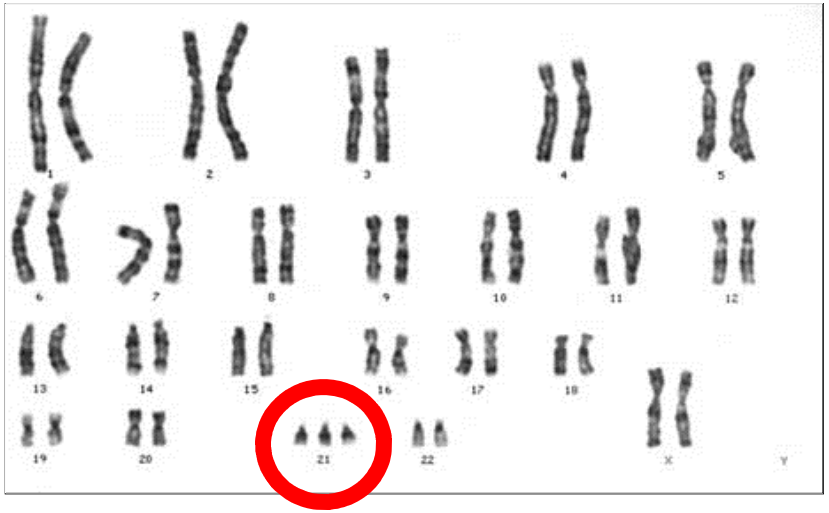


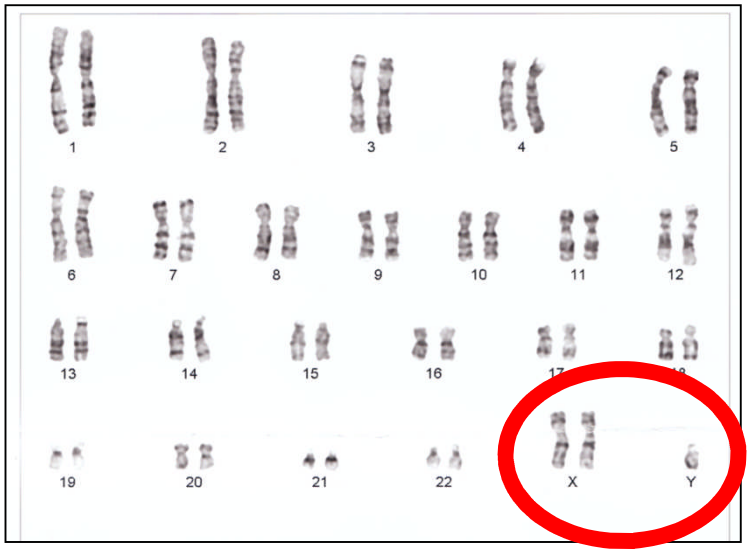
Alterations in Chromosomal Number

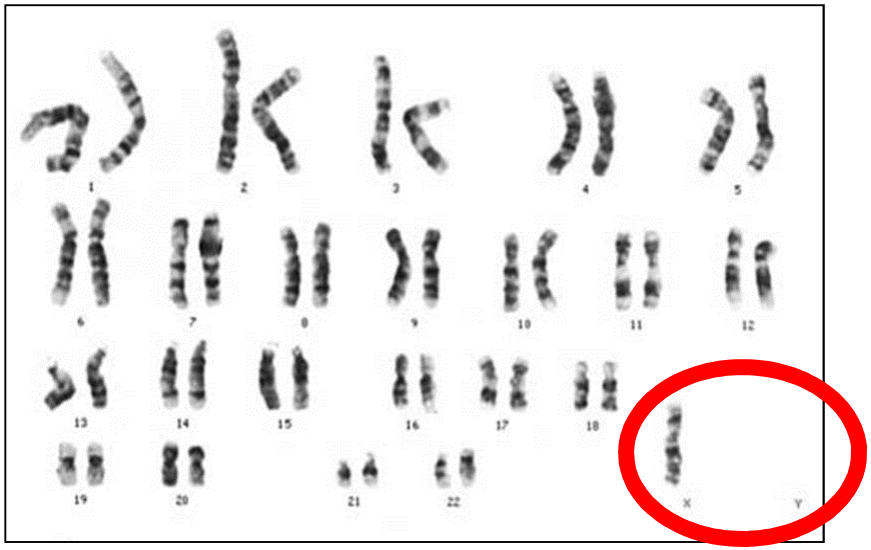
- Polyploidy:
  - Individuals have three copies of the autosomes, plus an extra sex chromosome (karyotype: 69, XXX, 69 XXY, 69 XYY)
- Aneuploidy:
  - One or more chromosomes are lacking or present in excess
    Examples: Trisomy 21 (Down's Syndrome), Klinefelter's syndrome (47, XXY), and Turner's syndrome (45, X)

Chromosomal Abnormalities







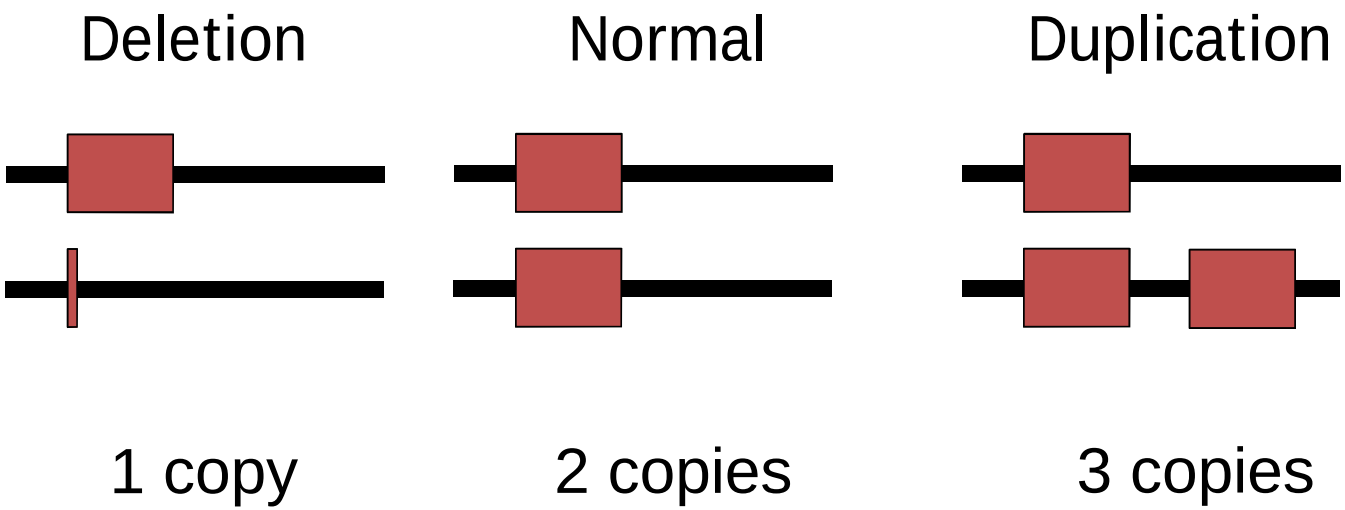


Klinefelter's Syndrome

Turner's Syndrome

Copy Number Variation

Chromosomal micro-duplications and deletions



CNV in schizophrenia

www.sciencemag.org SCIENCE VOL 320 25 APRIL 2008

Rare Structural Variants Disrupt Multiple Genes in Neurodevelopmental Pathways in Schizophrenia

Tom Walsh,[1*] Jon M. McClellan,[2*†] Shane E. McCarthy,[3*] Anjené M. Addington,[4*] Sarah B. Pierce,[1] Greg M. Cooper,[5] Alex S. Nord,[5] Mary Kusenda,[3,6] Dheeraj Malhotra,[3] Abhishek Bhandari,[3] Sunday M. Stray,[1] Caitlin F. Rippey,[5] Patricia Roccanova,[3] Vlad Makarov,[3] B. Lakshmi,[3] Robert L. Findling,[7] Linmarie Sikich,[8] Thomas Stromberg,[4] Barry Merriman,[9] Nitin Gogtay,[4] Philip Butler,[4] Kristen Eckstrand,[4] Laila Noory,[4] Peter Gochman,[4] Robert Long,[4] Zugen Chen,[9] Sean Davis,[10] Carl Baker,[5] Evan E. Eichler,[5] Paul S. Meltzer,[10] Stanley F. Nelson,[9] Andrew B. Singleton,[11] Ming K. Lee,[1] Judith L. Rapoport,[4] Mary-Claire King,[1,5] Jonathan Sebat[3]

nature Vol 455|11 September 2008|doi:10.1038/nature07229

LETTERS

Large recurrent microdeletions associated with schizophrenia

Hreinn Stefansson[1]*, Dan Rujescu[2]*, Sven Cichon[3,4]*, Olli P. H. Pietiläinen[5], Andres Ingason[1], Stacy Steinberg[1], Ragnheidur Fossdal[1], Engilbert Sigurdsson[6], Thordur Sigmundsson[6], Jacobine E. Buizer-Voskamp[7], Thomas Hansen[8,9], Klaus D. Jakobsen[8,9], Pierandrea Muglia[10], Clyde Francks[10], Paul M. Matthews[11], Arnaldur Gylfason[1], Bjarni V. Halldorsson[1], Daniel Gudbjartsson[1], Thorgeir E. Thorgeirsson[1], Asgeir Sigurdsson[1], Adalbjorg Jonasdottir[1], Aslaug Jonasdottir[1], Asgeir Bjornsson[1], Sigurborg Mattiasdottir[1], Thorarinn Blondal[1], Magnus Haraldsson[6], Brynja B. Magnusdottir[6], Ina Giegling[2], Hans-Jürgen Möller[2], Annette Hartmann[2], Kevin V. Shianna[12], Dongliang Ge[12], Anna C. Need[12], Caroline Crombie[13], Gillian Fraser[13], Nicholas Walker[14], Jouko Lonnqvist[15], Jaana Suvisaari[15], Annamarie Tuulio-Henriksson[15], Tiina Paunio[5,15], Timi Toulopoulou[16], Elvira Bramon[16], Marta Di Forti[16], Robin Murray[16], Mirella Ruggeri[17], Evangelos Vassos[16], Sarah Tosato[17], Muriel Walshe[16], Tao Li[16,18], Catalina Vasilescu[3], Thomas W. Mühleisen[3], August G. Wang[19], Henrik Ullum[20], Srdjan Djurovic[21,22], Ingrid Melle[22], Jes Olesen[23], Lambertus A. Kiemeney[24], Barbara Franke[25], GROUP†, Chiara Sabatti[26], Nelson B. Freimer[27], Jeffrey R. Gulcher[1], Unnur Thorsteinsdottir[1], Augustine Kong[1], Ole A. Andreassen[21,22], Roel A. Ophoff[7,27], Alexander Georgi[28], Marcella Rietschel[28], Thomas Werge[8], dstein[12], Markus M. Nöthen[3,4], Leena Peltonen[5,29,30], David A. Collier[16,18], David St

doi:10.1038/nature07239 nature

Rare chromosomal deletions and duplications increase risk of schizophrenia

The International Schizophrenia Consortium*

Human Molecular Genetics, 2009, Vol. 18, No. 8 **1497–1503**
doi:10.1093/hmg/ddp043
Advance Access published on January 29, 2009

Support for the involvement of large copy number variants in the pathogenesis of schizophrenia

George Kirov, Detelina Grozeva, Nadine Norton, Dobril Ivanov, Kiran K. Mantripragada, Peter Holmans, International Schizophrenia Consortium†, the Wellcome Trust Case Control Consortium†, Nick Craddock, Michael J. Owen* and Michael C. O'Donovan

CNV in autism

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 FEBRUARY 14, 2008 VOL. 358 NO. 7

Association between Microdeletion and Microduplication at 16p11.2 and Autism

Lauren A. Weiss, Ph.D., Yiping Shen, Ph.D., Joshua M. Korn, B.S., Dan E. Arking, Ph.D., David T. Miller, M.D., Ph.D., Ragnheidur Fossdal, B.Sc., Evald Saemundsen, B.A., Hreinn Stefansson, Ph.D., Manuel A.R. Ferreira, Ph.D., Todd Green, B.S., Orah S. Platt, M.D., Douglas M. Ruderfer, M.S., Christopher A. Walsh, M.D., Ph.D., David Altshuler, M.D., Ph.D., Aravinda Chakravarti, Ph.D., Rudolph E. Tanzi, Ph.D., Kari Stefansson, M.D., Ph.D., Susan L. Santangelo, Sc.D., James F. Gusella, Ph.D., Pamela Sklar, M.D., Ph.D., Bai-Lin Wu, M.Med., Ph.D., and Mark J. Daly, Ph.D., for the Autism Consortium

Strong Association of De Novo Copy Number Mutations with Autism

Jonathan Sebat,[1*] B. Lakshmi,[1] Dheeraj Malhotra,[1*] Jennifer Troge,[1*] Christa Lese-Martin,[2] Tom Walsh,[3] Boris Yamrom,[1] Seungtai Yoon,[1] Alex Krasnitz,[1] Jude Kendall,[1] Anthony Leotta,[1] Deepa Pai,[1] Ray Zhang,[1] Yoon-Ha Lee,[1] James Hicks,[1] Sarah J. Spence,[4] Annette T. Lee,[5] Kaija Puura,[6] Terho Lehtimäki,[7] David Ledbetter,[2] Peter K. Gregersen,[5] Joel Bregman,[8] James S. Sutcliffe,[9] Vaidehi Jobanputra,[10] Wendy Chung,[10] Dorothy Warburton,[10] Mary-Claire King,[3] David Skuse,[11] Daniel H. Geschwind,[12] T. Conrad Gilliam,[13] Kenny Ye,[14] Michael Wigler[1]†

www.sciencemag.org SCIENCE VOL 316 20 APRIL 2007

Vol 459|28 May 2009|doi:10.1038/nature07953 nature

LETTERS

Structural Variation of Chromosomes in Autism Spectrum Disorder

Christian R. Marshall,[1] Abdul Noor,[2] John B. Vincent,[2] Anath C. Lionel,[1] Lars Mary Shago,[3] Rainald Moessner,[1] Dalila Pinto,[1] Yan Ren,[1] Bhooma Thiruvahin Andreas Fiebig,[6] Stefan Schreiber,[6] Jan Friedman,[7] Cees E.J. Ketelaars,[8] Yvonne Susan Kirkpatrick,[10] Rob Nicolson,[11] Leon Sloman,[2] Anne Summers,[12] Clare A. Ahmad Teebi,[4] David Chitayat,[4] Rosanna Weksberg,[4] Ann Thompson,[13] Cathy Vicki Crosbie,[14] Sandra Luscombe,[14] Rebecca Baatjes,[1] Lonnie Zwaigenbaum,[15] Bridget Fernandez,[14] Peter Szatmari,[13] and Stephen W. Scherer[1,*]

The American Journal of Human Genetics *82*, 477–488, February 2008

Autism genome-wide copy number variation reveals ubiquitin and neuronal genes

Joseph T. Glessner[1], Kai Wang[1], Guiqing Cai[2], Olena Korvatska[3], Cecilia E. Kim[1], Shawn Wood[4], Haitao Zhang[1], Annette Estes[3], Camille W. Brune[5], Jonathan P. Bradfield[1], Marcin Imielinski[1], Edward C. Frackelton[1], Jennifer Reichert[2], Emily L. Crawford[6], Jeffrey Munson[3], Patrick M. A. Sleiman[1], Rosetta Chiavacci[1], Kiran Annaiah[1], Kelly Thomas[1], Cuiping Hou[1], Wendy Glaberson[1], James Flory[1], Frederick Otieno[1], Maria Garris[1], Latha Soorya[2], Lambertus Klei[4], Joseph Piven[7], Kacie J. Meyer[8], Evdokia Anagnostou[2], Takeshi Sakurai[2], Rachel M. Game[6], Danielle S. Rudd[8], Danielle Zurawiecki[2], Christopher J. McDougle[10], Lea K. Davis[8], Judith Miller[9], David J. Posey[10], Shana Michaels[4], Alexander Kolevzon[2], Jeremy M. Silverman[2], Raphael Bernier[3], Susan E. Levy[11], Robert T. Schultz[11], Geraldine Dawson[3], Thomas Owley[5], William M. McMahon[9], Thomas H. Wassink[8], John A. Sweeney[5], John I. Nurnberger Jr[10], Hilary Coon[9], James S. Sutcliffe[6], Nancy J. Minshew[12], Struan F. A. Grant[1,11], Maja Bucan[13], Edwin H. Cook Jr[5], Joseph D. Buxbaum[2,14], Bernie Devlin[4], Gerard D. Schellenberg[15] & Hakon Hakonarson[1,11]

CNV in schizophrenia

Table 2 Combined results of previous studies and the current data-set[a]

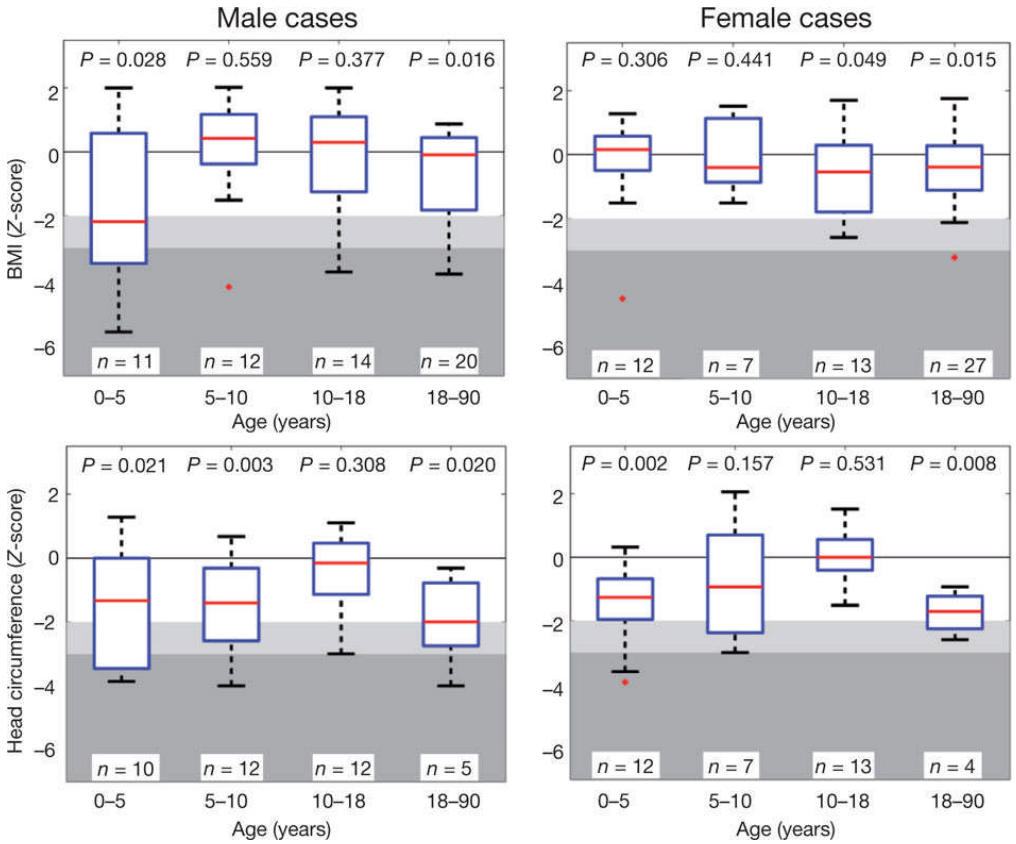
Locus	*P*-value in previous studies	CNV frequency, % (*n*/*N*)		OR (95% CI)	*P*
		Case group	Control group		
1q21.1 del	1.3×10^{-9}	0.17 (33/19056)	0.021 (17/81821)	8.35 (4.65–14.99)	4.1×10^{-13}
1q21.1 dup	2.0×10^{-4}	0.13 (21/16247)	0.037 (24/64046)	3.45 (1.92–6.20)	9.9×10^{-5}
NRXN del	7.9×10^{-9}	0.18 (33/18762)	0.020 (10/51161)	9.01 (4.44–18.29)	1.3×10^{-11}
3q29 del	2.3×10^{-8}	0.082 (14/17005)	0.0014 (1/69965)	57.65 (7.58–438.44)	1.5×10^{-9}
WBS dup	5.5×10^{-5}	0.066 (14/21269)	0.0058 (2/34455)	11.35 (2.58–49.93)	6.9×10^{-5}
VIPR2 dup	0.006	0.11 (15/14218)	0.069 (17/24815)	1.54 (0.77–3.09)	0.27
15q11.2 del	2.2×10^{-7}	0.59 (116/19547)	0.28 (227/81802)	2.15 (1.71–2.68)	2.5×10^{-10}
AS/PWS dup	0.014	0.083 (12/14464)	0.0063 (3/47686)	13.20 (3.72–46.77)	5.6×10^{-6}
15q13.3 del	2.1×10^{-11}	0.14 (26/18571)	0.019 (15/80422)	7.52 (3.98–14.19)	4.0×10^{-10}
16p13.11 dup	0.03	0.31 (37/12029)	0.13 (93/69289)	2.30 (1.57–3.36)	5.7×10^{-5}
16p11.2 distal del	0.0014	0.063 (13/20732)	0.018 (5/27045)	3.39 (1.21–9.52)	0.017
16p11.2 dup	3.2×10^{-14}	0.35 (58/16772)	0.030 (19/63068)	11.52 (6.86–19.34)	2.9×10^{-24}
17p12 del	0.0004	0.094 (12/12773)	0.026 (17/65402)	3.62 (1.73–7.57)	0.0012
17q12 del	0.004	0.036 (5/14024)	0.0054 (4/74447)	6.64 (1.78–24.72)	0.0072
22q11.2 del	1.0×10^{-30}	0.29 (56/19084)	0.00 (0/77055)	NA (28.27–∞)	4.4×10^{-40}

del, deletion; dup, duplications; NA, not applicable; WBS, Williams–Beuren syndrome; AS/PWS, Angelman/Prader–Willi syndrome.
a. For a more detailed version of this table that includes the CNV frequency, % (*n*/*N*) from previous studies see online Table DS6. *P*-values are based on Fisher exact test, 2-tailed.

Rees et al., British Journal of Psychiatry 2014, 204 (2) 108-114

16p11.2 CNV region

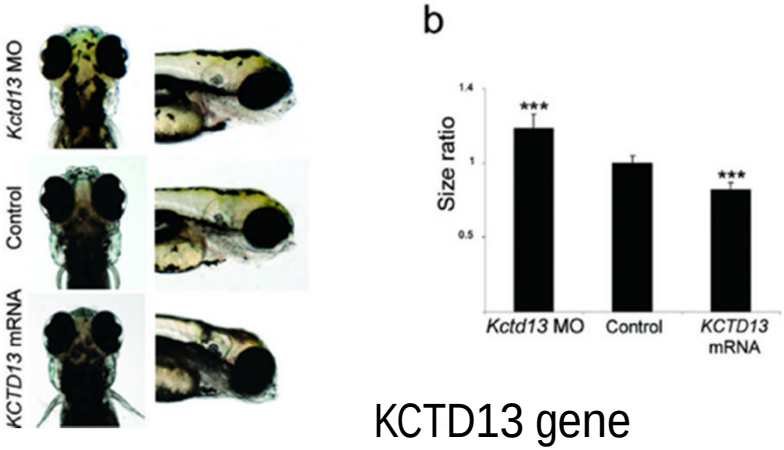
Effect of the chromosome 16p11.2 duplication on BMI and head circumference.



Jacquemont et al. Nature (2011) 478, 97–102
doi:10.1038/nature10406

16p11.2 Duplications and Deletions produce mirror phenotypes on obesity and many measures of brain imaging

Similar phenotypes are seen in a zebrafish model using the KCTD13 gene



KCTD13 gene

Golzio et al Nature. (2012) 485:363-7

Many phenotype associated loci effect non-coding genomic regions

- There are important non-coding regions that regulate gene expression (e.g., promoters, UTRs), and regions of the genome that are transcribed but not translated
- May alter epigenetic processes and the structure of chromatin?

Learning Objectives

Understand:

- The distinction between Quantitative Genetic and Molecular Genetic research
- The structure of DNA
- How DNA functions as the hereditary material
  - How DNA is packaged
  - How DNA function is regulated, etc
- The different types of genetic diversity in human populations
- Common terms used
- The broad types of genotype-phenotype relationships

Online Resources

- Help Me Understand Genetics – ebook (free) primer on genetics
- https://ghr.nlm.nih.gov/primer

- Your Genome: Resources on DNA, Genomes and Proteins
- http://www.yourgenome.org/

- DNA to Protein
- http://www.yourgenome.org/video/from-dna-to-protein

- DNA Replication
- http://www.yourgenome.org/video/dna-replication

- DNA Sequencing
- http://www.yourgenome.org/video/dna-sequencing

- Also: Genetic Science Learning Centre
- http://learn.genetics.utah.edu/

Learning Questions

- Why might a genetic variant in a coding region have no impact on a protein?
- How would you determine whether a gene is expressed in the brain?
- Describe an experiment that would allow you to determine whether common genetic variants were associated with a trait of interest.