equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, 40, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc.*, Geophys. Supp., 5, 285 (1949).
[3] Von Arx, W. S., *Woods Hole Papers in Phys. Oceanog. Meteor.*, 11 (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, 2 (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3',5' linkages. The two chains (but 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 is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

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

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

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

King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK

Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, 171, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, 39, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, 6, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, 9, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, 36, 201 (1952).
[5] Astbury, W. T., *Symp. Soc. Exp. Biol.* 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, 10, 192 (1953).

Molecular Structure of Deoxypentose Nucleic Acids

While the biological properties of deoxypentose nucleic acid suggest a molecular structure containing great complexity, X-ray diffraction studies described here (cf. Astbury[1]) show the basic molecular configuration has great simplicity. The purpose of this communication is to describe, in a preliminary way, some of the experimental evidence for the polynucleotide chain configuration being helical, and existing in this form when in the natural state. A fuller account of the work will be published shortly.

The structure of deoxypentose nucleic acid is the same in all species (although the nitrogen base ratios alter considerably) in nucleoprotein, extracted or in cells, and in purified nucleate. The same linear group of polynucleotide chains may pack together parallel in different ways to give crystalline[1-3], semi-crystalline or paracrystalline material. In all cases the X-ray diffraction photograph consists of two regions, one determined largely by the regular spacing of nucleotides along the chain, and the other by the longer spacings of the chain configuration. The sequence of different nitrogen bases along the chain is not made visible.

Oriented paracrystalline deoxypentose nucleic acid ('structure B' in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-A. reflexion corresponded to the internucleotide repeat along the fibre axis. The ~34 A. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the interstitial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown[5] (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the *n*th layer line being proportional to the square of J_n, the *n*th order Bessel function. A straight line may be drawn approximately through the innermost maxima of each Bessel function and the origin. The angle this line makes with the equator is roughly equal to the angle between an element of the helix and the helix axis. If a unit repeats n times along the helix there will be a meridional reflexion (J_0^2) on the *n*th layer line. The helical configuration produces side-bands on this fundamental frequency, the effect being to reproduce the intensity distribution about the origin around the new origin, on the *n*th layer line, corresponding to C in Fig. 2.

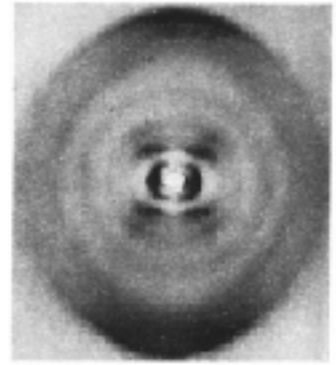


Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

We will now briefly analyse in physical terms some of the effects of the shape and size of the repeat unit or nucleotide on the diffraction pattern. First, if the nucleotide consists of a unit having circular symmetry about an axis parallel to the helix axis, the whole diffraction pattern is modified by the form factor of the nucleotide. Second, if the nucleotide consists of a series of points on a radius at right-angles to the helix axis, the phases of radiation scattered by the helices of different diameter passing through each point are the same. Summation of the corresponding Bessel functions gives reinforcement for the inner-

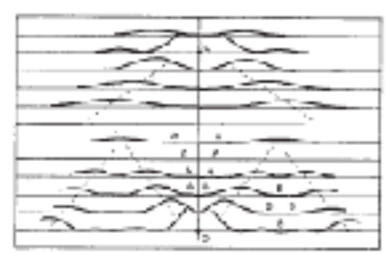
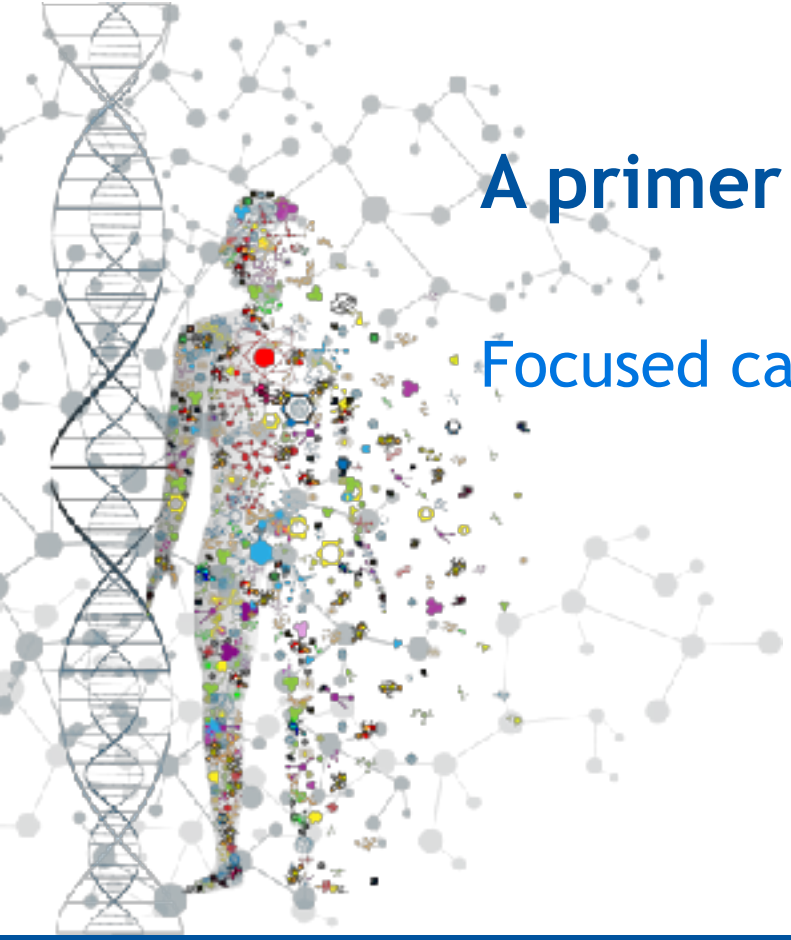


Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid. The squares of Bessel functions are plotted about 0 on the equator and on the first, second, third and fifth layer lines for half of the nucleotide mass at 20 A. diameter and remainder distributed along a radius, the mass at a given radius being proportional to the radius. About C on the tenth layer line similar functions are plotted for an outer diameter of 12 A.



A primer in complex human genetics

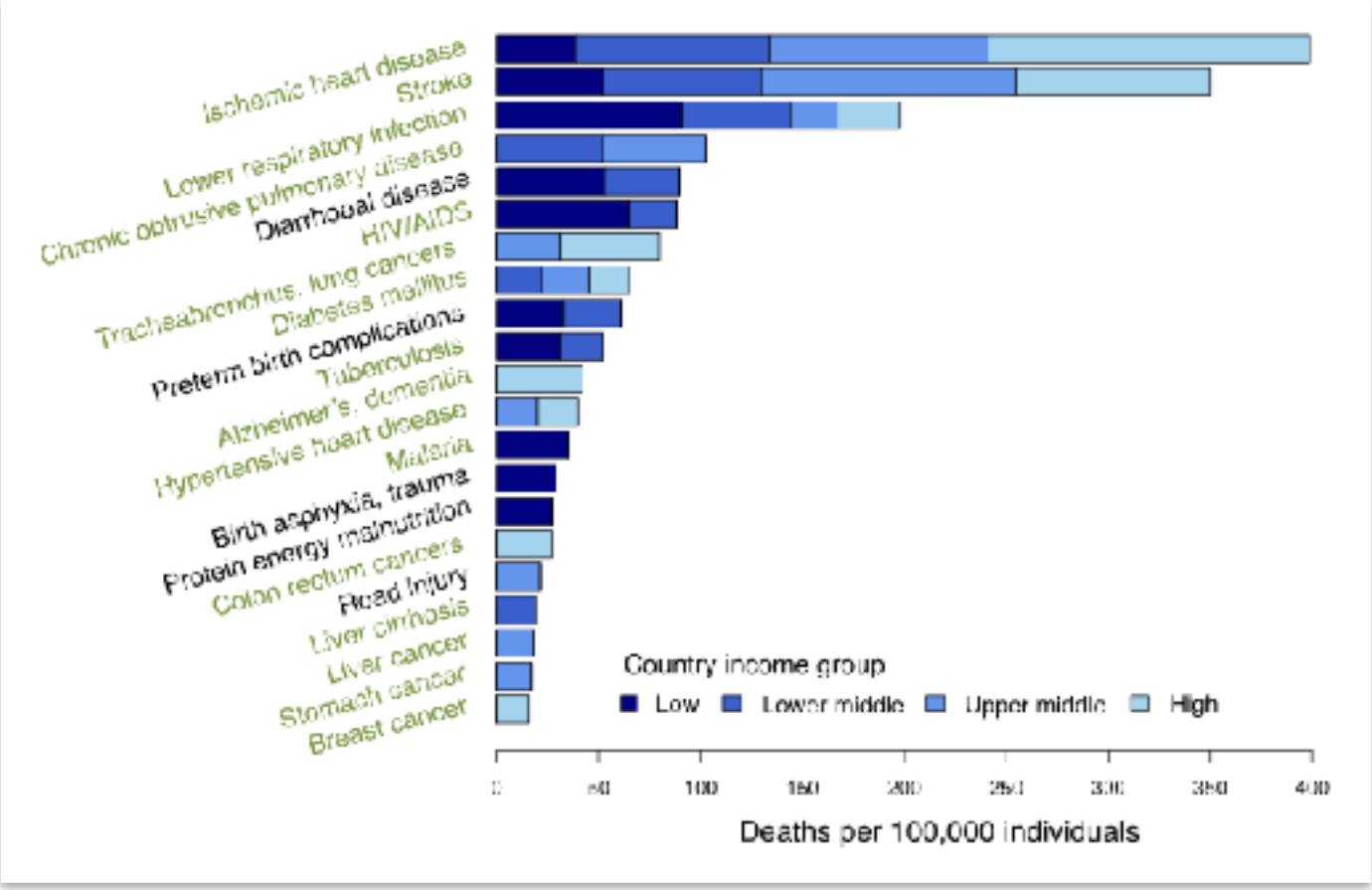
Focused cardiovascular disease

Sander W. van der Laan, PhD
s.w.vanderlaan-2@umcutrecht.nl • @swvanderlaan • swvanderlaan.github.io

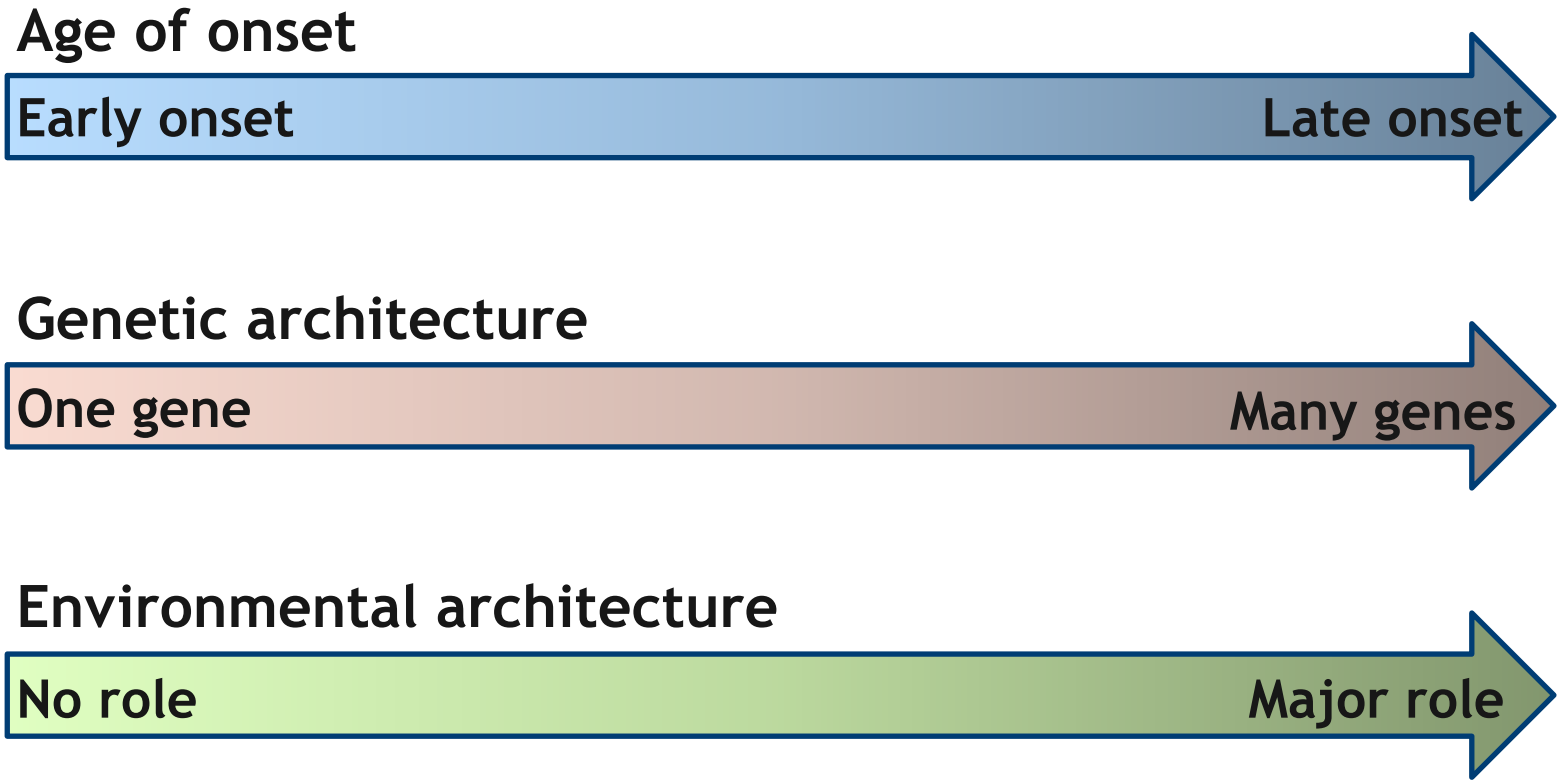
UMC Utrecht
Center for Circulatory Health

UMC Utrecht
Athero-Express Biobank Studies

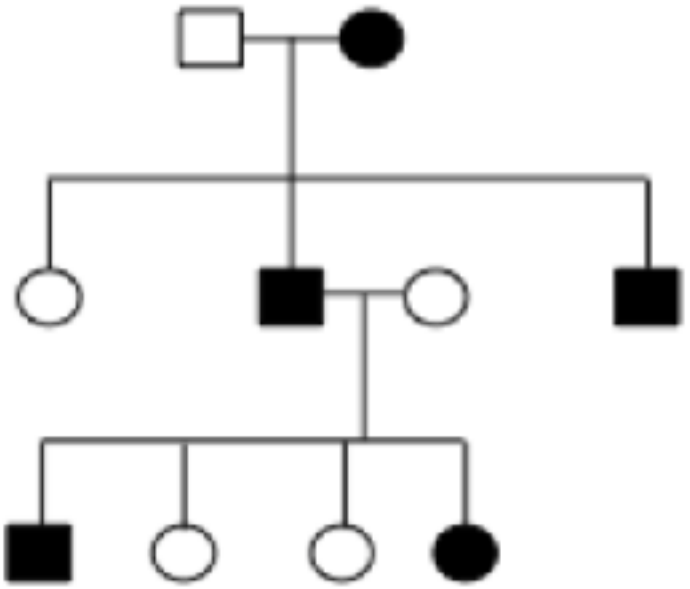
Human disease around the globe



The spectrum(s) of disease



Rare (Mendelian) diseases



family-based studies

Genotype

Disease

Environment

Rare (Mendelian) diseases

1983

A polymorphic DNA marker genetically linked to Huntington's disease

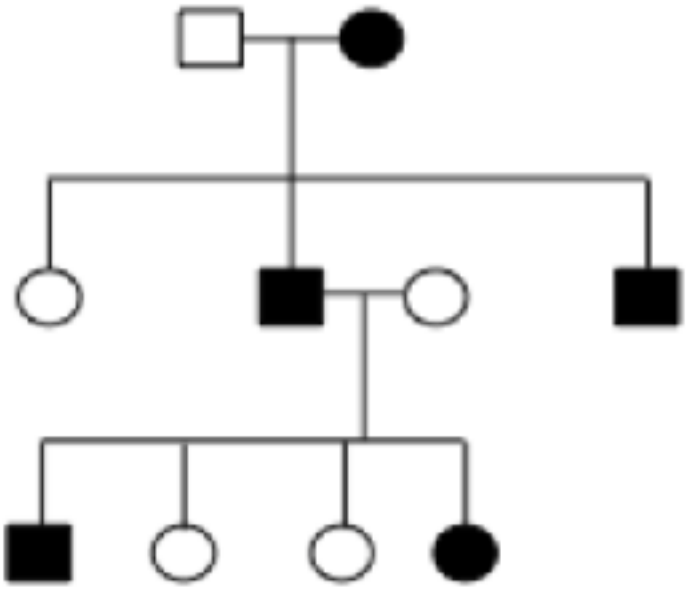
James F. Gusella*, Nancy S. Wexler†‖, P. Michael Conneally‡, Susan L. Naylor§
Mary Anne Anderson*, Rudolph E. Tanzi*, Paul C. Watkins*¶, Kathie
Margaret R. Wallace†, Alan Y. Sakaguchi§, Anne B. Young‖, Ira Sh
Ernesto Bonilla‖ & Joseph B. Martin*

* Neurology Department and Genetics Unit, Massachusetts General Hospital and Harvard Medical School, Boston, M
† Hereditary Disease Foundation, 9701 Wilshire Blvd, Beverley Hills, California 90212, US
‡ Department of Medical Genetics, Indiana University Medical Center, Indianapolis, Indiana 4624
§ Department of Human Genetics, Roswell Park Memorial Institute, Buffalo, New York 14263
‖ Venezuela Collaborative Huntington's Disease Project

Family studies show that the Huntington's disease gene is linked to a polymorphic DNA marke chromosome 4. The chromosomal localization of the Huntington's disease gene is the first step DNA technology to identify the primary genetic defect in this disorder.

1989

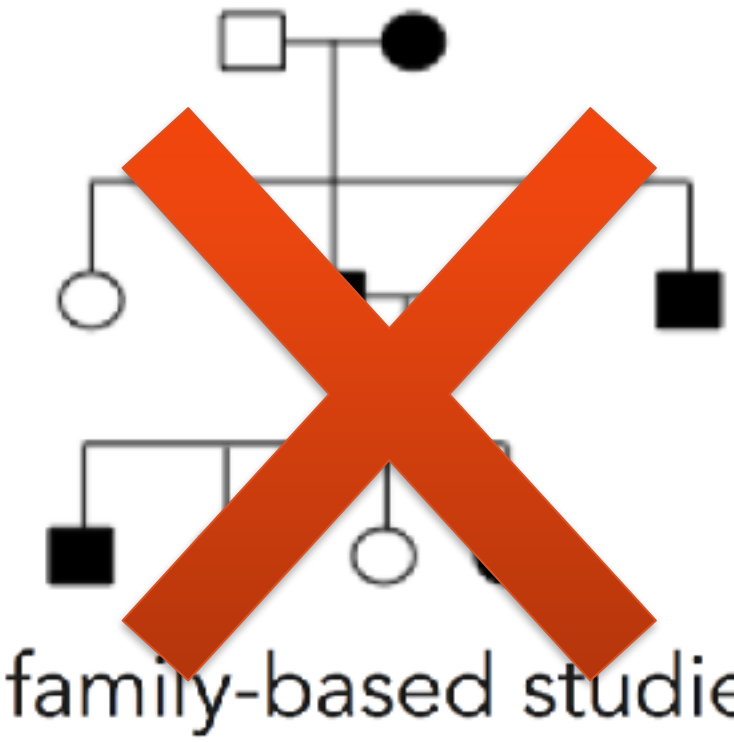
Common (complex) diseases

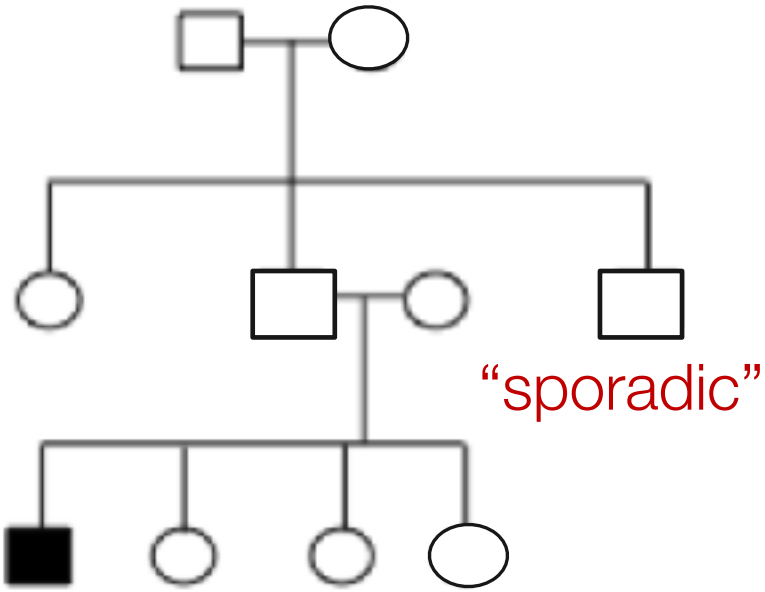


family-based studies

?

Common (complex) diseases





The challenges of common disease

- Heterogeneity
- Late (or broad age range for) onset
- Interaction of genes and environment (multifactorial)
- Overlap with other diseases

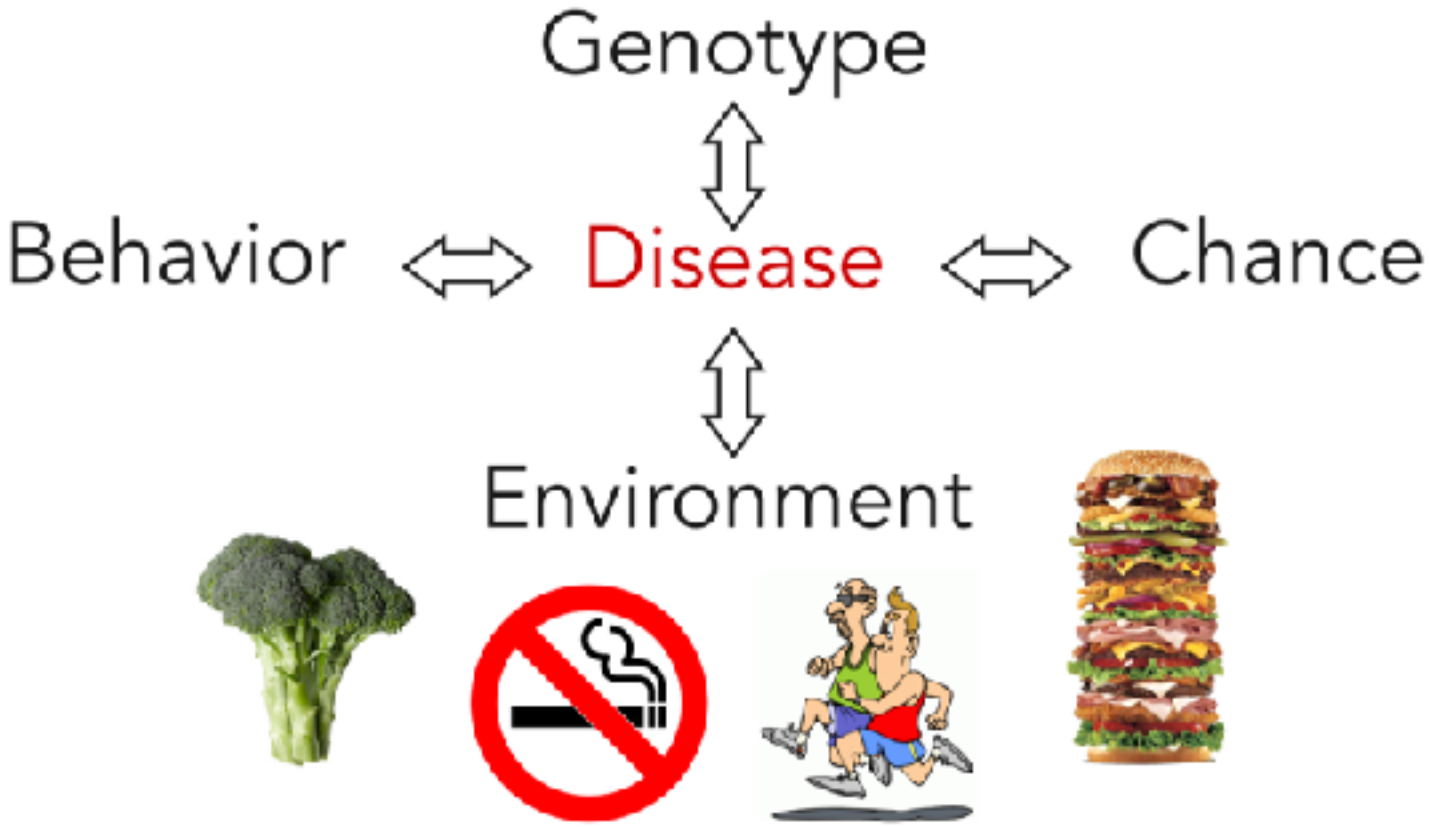
For the Elderly, Diseases That Overlap

1 2 3 4 5 6 7 NEXT »

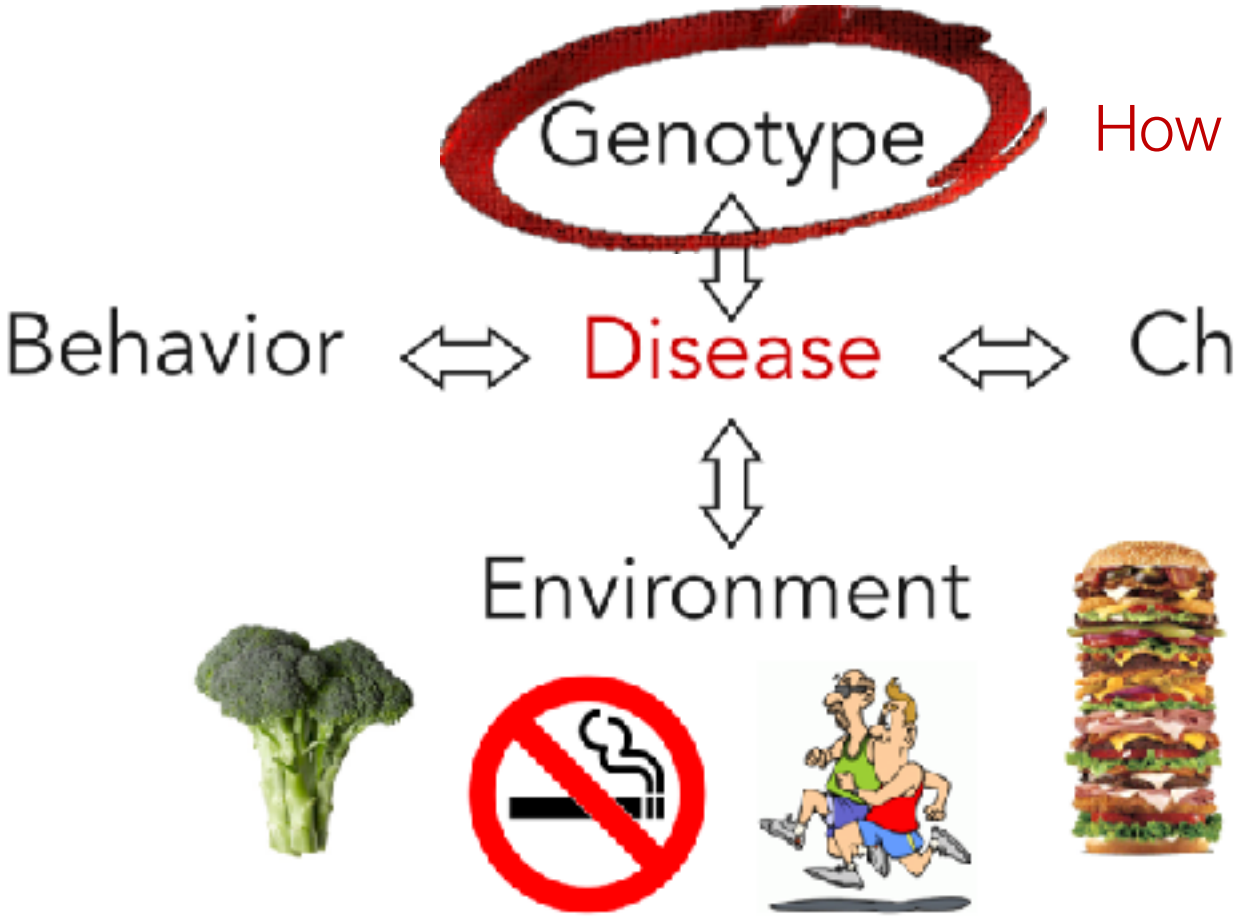
Researchers are beginning to focus more intently on the overlaps and possible interconnections, and some scientists argue that it may not be possible to treat dementia without treating vascular problems.

The New York Times, 15 April 2013

Multifactorial disease



Multifactorial disease



How do we know this matters?

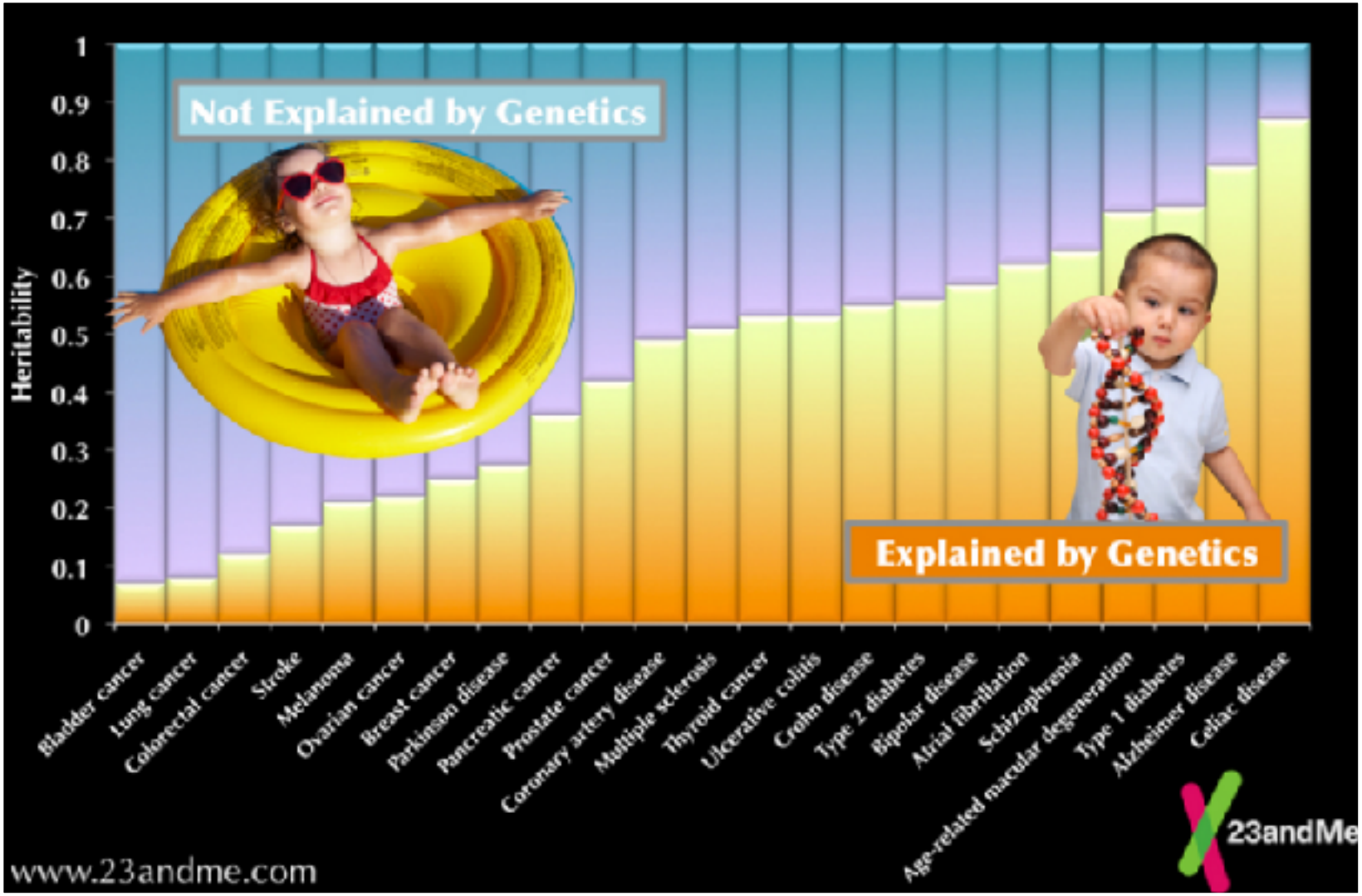
Heritability

Given I am a patient, what is risk of disease for…

	Type 1	Type 2
Your neighbor (unrelated)?	0.4%	5-10%
Your sibling?	6%	30%
Your identical twin?	30-50%	>80%

PIW de Bakker

The range of heritability estimates



Family history

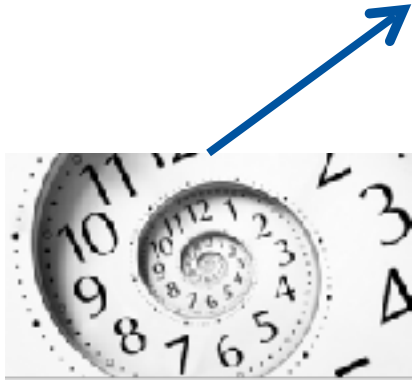
- Framingham Heart Study | www.framinghamheartstudy.org
  - A positive history of cardiovascular disease and associated risk factors tend to aggregate in families
  - Familial aggregation heritability of CVD estimated ≥90% (before 46 years)
  - Family history is an independent risk factor (FHS)
  - Positive family history associated with pre-clinical atherosclerosis as measured by carotid IMT, $h^2 \approx 0.35$
- High concordance rate among monozygotic twins, compared to dizygotic twins
- Heritability of atherosclerosis (carotid IMT) $h^2 \approx 0.21\text{-}0.64$ and is increased by age and cardiovascular risk factors

There is clearly a heritability factor for atherosclerotic and consequent cardiovascular disease

Why do some individuals have a higher risk for a disease than others?

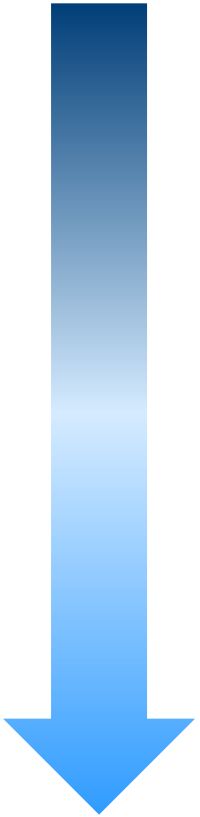
How can we alleviate disease burden in the human population?

Drug development



What's the goal of genetics?

- Understanding true causal disease pathways
  - Identify risk factors
  - Inform novel research directions
  - Enable rational and efficient drug development
- Precision medicine
  - Evaluate individual disease risk
  - Early disease identification or prevention
  - Understand patient's therapeutic response

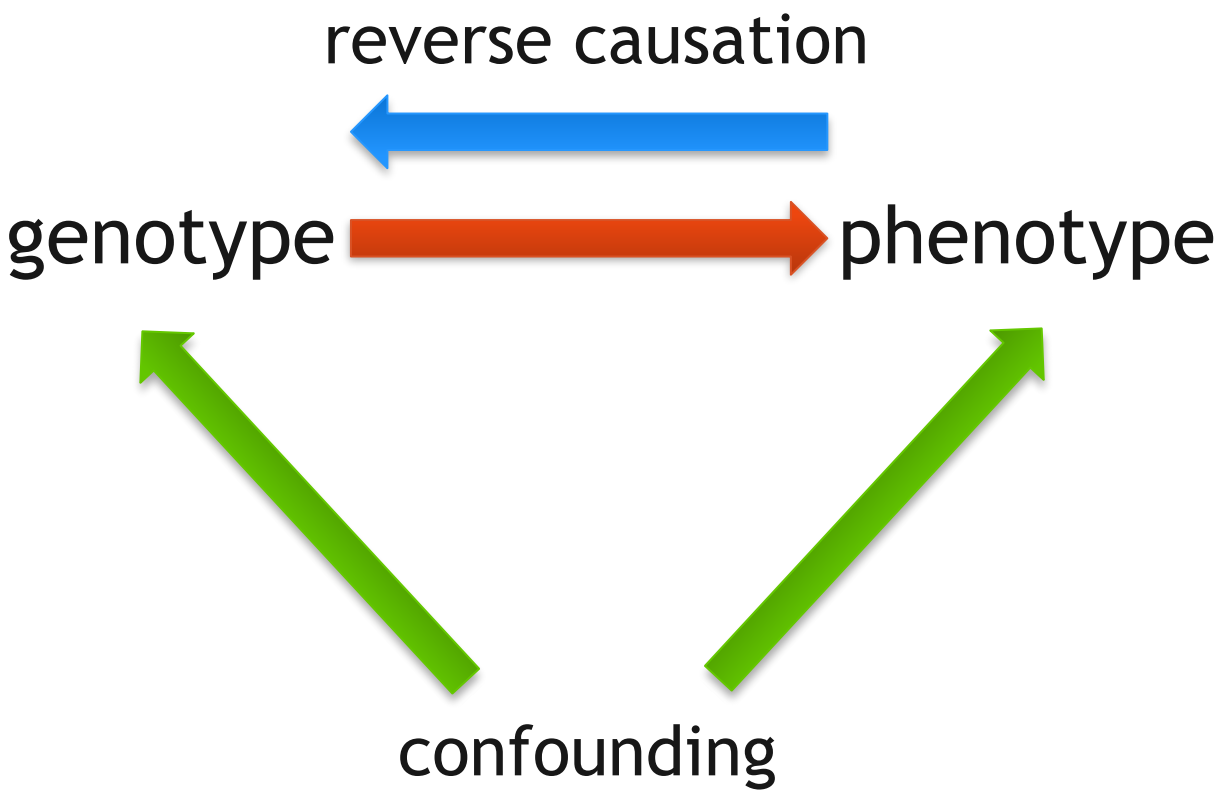


Why genetics at all?

- Genotypes are randomly assigned at meiosis
  - Nature's randomized clinical trial
- Genotypes are fixed and unaltered by the disease
  - Exception: somatic mutations in cancer
- We have become increasingly good at measuring genotypes
  - Lots and lots of data

The limitations of genetics

The limitations of genetics



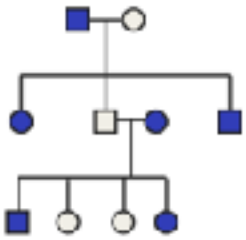
Where we've been and where we are

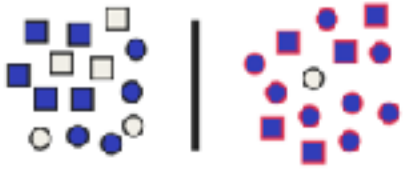
Linkage analysis

Candidate gene studies

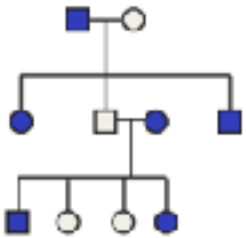
GWAS

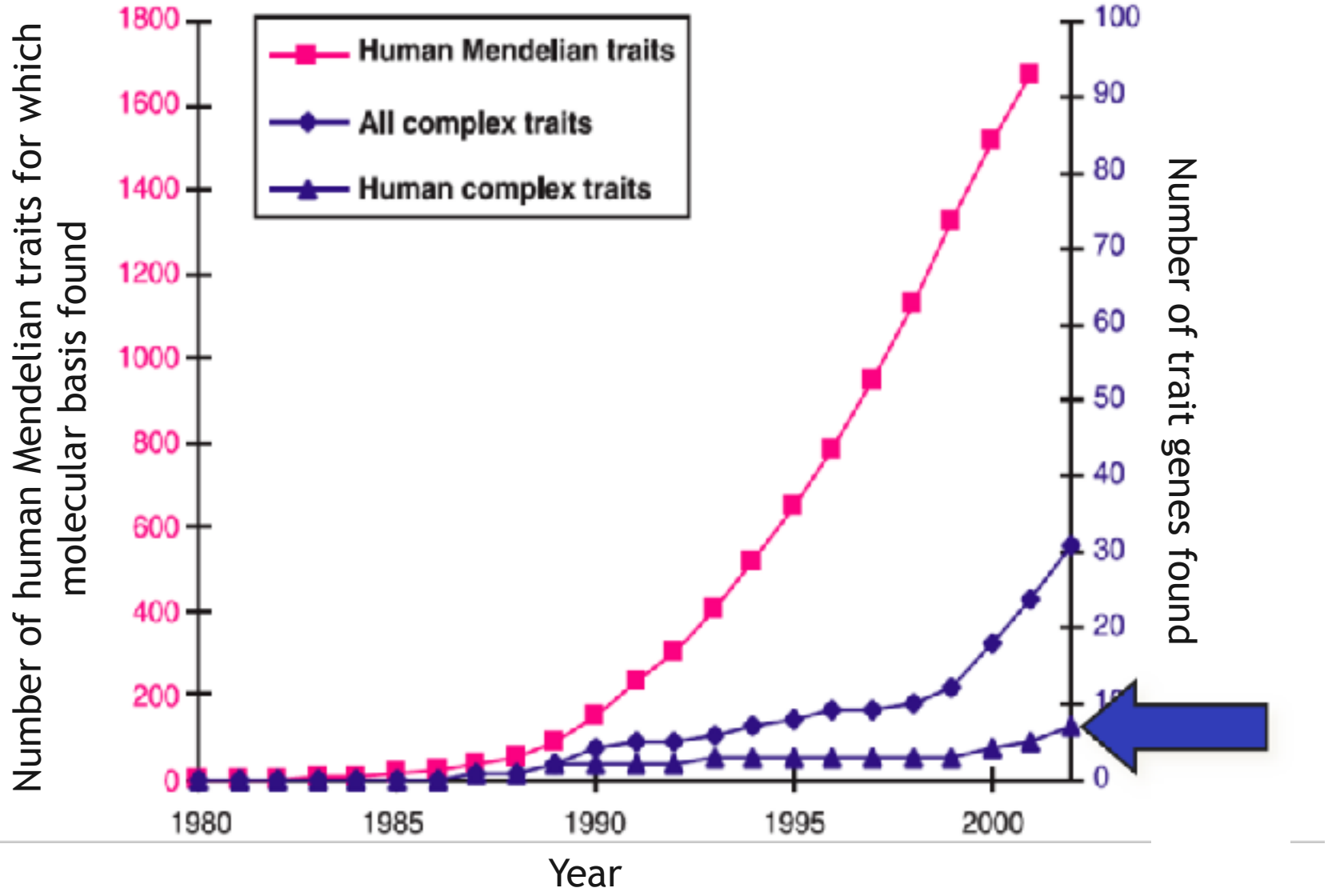
Sequencing





Linkage analysis

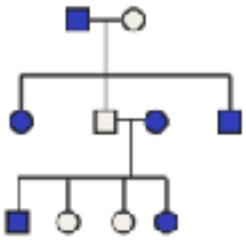


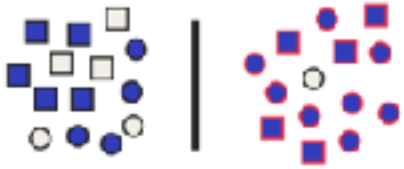


Glazier et al. Science 2002

Linkage analysis

Candidate gene studies





The candidate gene approach

- Pick a gene that might have a role in your disease

  arbitrary

- Genotype individuals at a few sites around that gene
  - Typically 1,000 – 2,000 samples

    no power

- Test genetic sites for association

A poor history of candidate gene studies

March/April 2002 · Vol. 4 · No. 2

review

A comprehensive review of genetic association studies

Joel N. Hirschhorn, MD, PhD[1–3], Kirk Lohmueller[1], Edward Byrne[1], and Kurt Hirschhorn, MD[4]

Most common diseases are complex genetic traits, with multiple genetic and environmental components contributing to susceptibility. It has been proposed that common genetic variants, including single nucleotide polymorphisms (SNPs), influence susceptibility to common disease. This proposal has begun to be tested in numerous studies of association between genetic variation at these common DNA polymorphisms and variation in disease susceptibility. We have performed an extensive review of such association studies. We find that over 600 positive associations between common gene variants and disease have been reported; these associations, if correct, would have tremendous importance for the prevention, prediction, and treatment of most common diseases. However, most reported associations are not robust: of the 166 putative associations which have been studied three or more times, only 6 have been consistently replicated. Interestingly, of the remaining 160 associations, well over half were observed again one or more times. We discuss the possible reasons for this irreproducibility and suggest guidelines for performing and interpreting genetic association studies. In particular, we emphasize the need for caution in drawing conclusions from a single report of an association between a genetic variant and disease susceptibility. **Genet Med 2002:4(2):45–61.**

Key Words: *human genetics, association studies, common disease, polymorphisms*

PloS Medicine, 2005

The candidate gene problem:

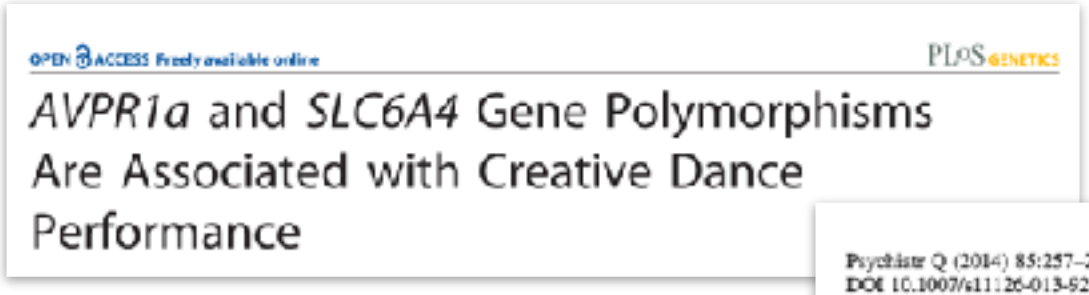
- Lack of statistical rigor
- Lack of large samples
- Lack of data quality control
- Lack of replication data

Need systematic, unbiased approach

Problems with the candidate gene approach

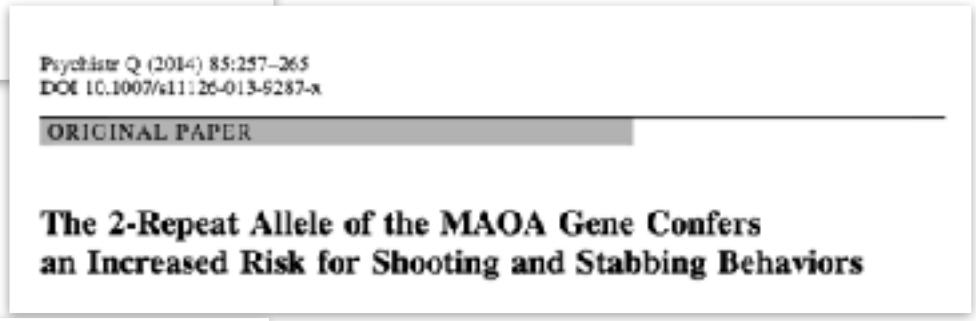
- Small sample sizes
- Weak effects
- No community-wide standards for QC, association claims
- Population stratification

Important side note: this still happens


OPEN ACCESS Freely available online

PLoS GENETICS

AVPR1a and *SLC6A4* Gene Polymorphisms Are Associated with Creative Dance Performance


Psychiatr Q (2014) 85:257–265
DOI 10.1007/s11126-013-9287-x

ORIGINAL PAPER

The 2-Repeat Allele of the MAOA Gene Confers an Increased Risk for Shooting and Stabbing Behaviors


SCIENTIFIC REPORTS

OPEN

The association between romantic relationship status and 5-HT1A gene in young adults

SUBJECT AREAS:
HUMAN BEHAVIOUR
BEHAVIOURAL GENETICS