The OMICs Revolution

Toni I. Gossmann

TU Dortmund University

March 2025

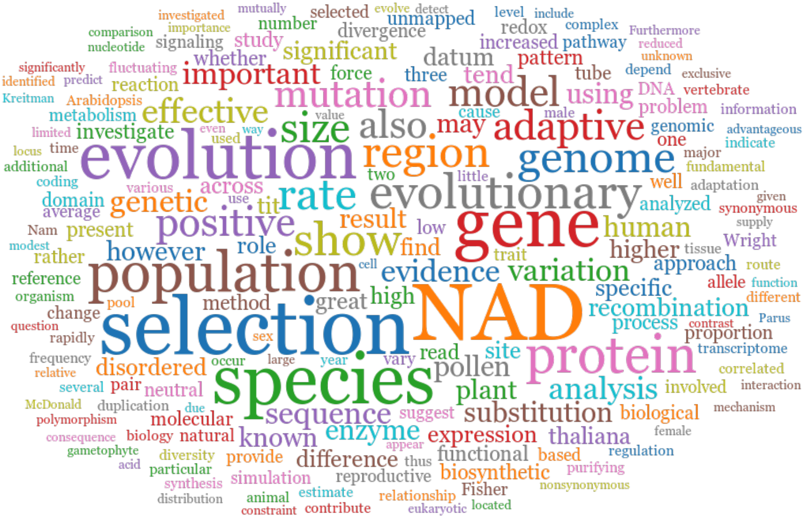
My group

Computational
Systems Biology

My group

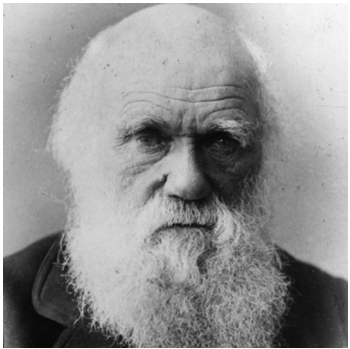
My group

My research as a word cloud



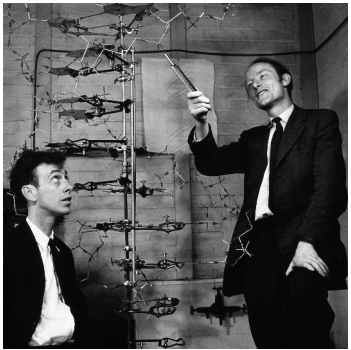
Evolution

What Darwin also realized



Natural selection requires

- Variation
- Heritability



On the molecular level?

- DNA differs between individuals
- Germline mutations can be inherited

Why is the right time to ask evolutionary questions?

Huge amounts of data to address biological/medical questions

We live in an **OMICS** (the totally of some sort) world

- Genomics, Transcriptomics, Proteomics, Metabolomics, Lipidomics, Meta-Genomics, Methylomics, ...

Rely mainly on two technologies:

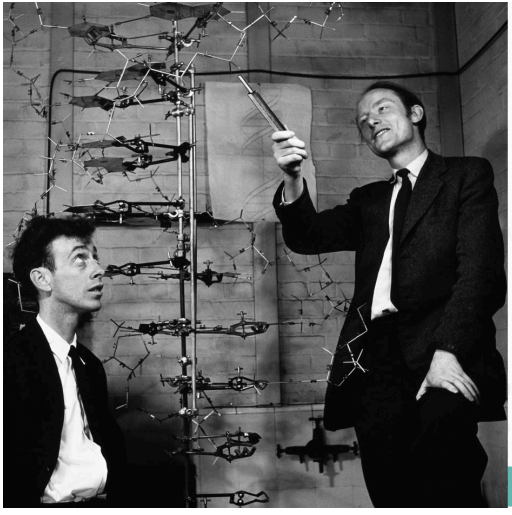
- Mass spectrometry
- **DNA sequencing**

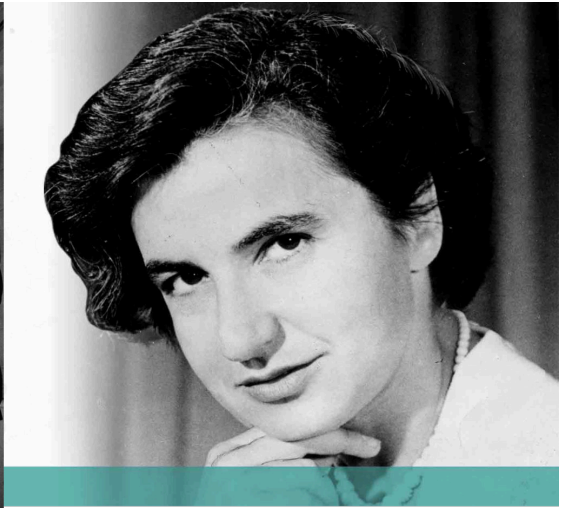
Hypotheses-driven versus **Data-driven**

Today's plan

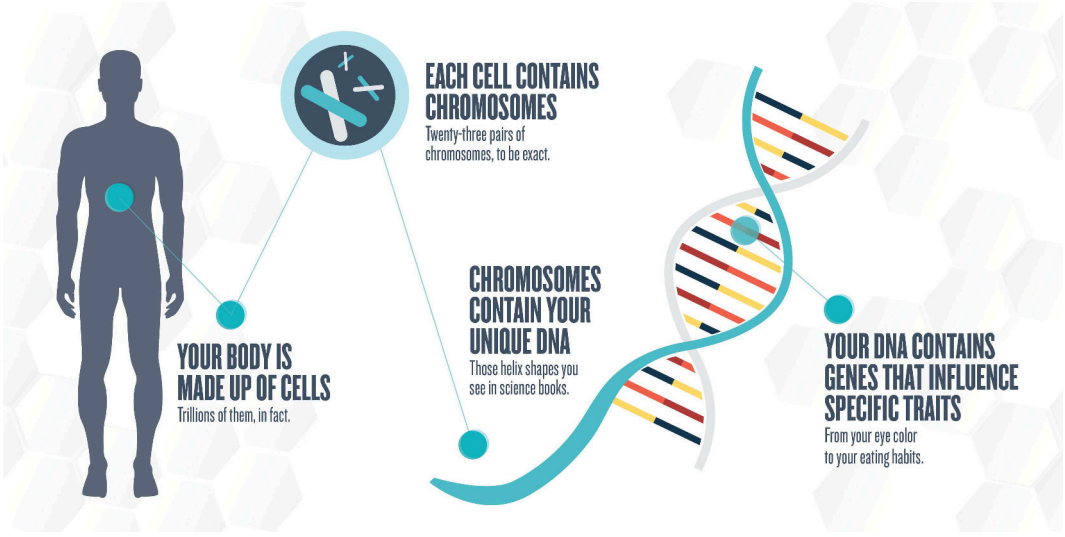
1. Introduction ("The basics")
2. Journal club
3. Own research example

The pioneers

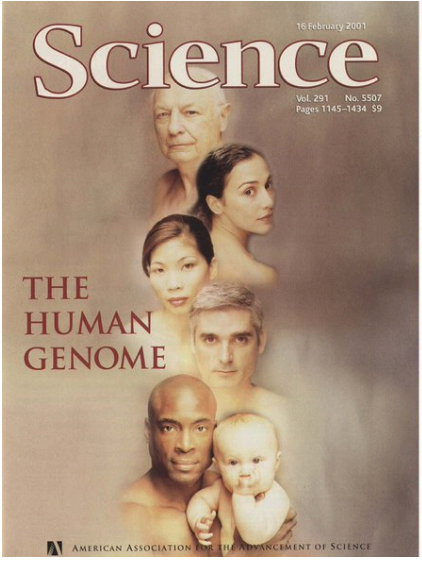


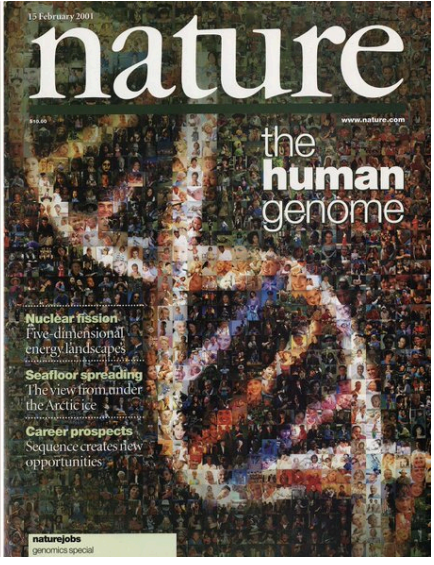


What is DNA?


EACH CELL CONTAINS CHROMOSOMES
Twenty-three pairs of chromosomes, to be exact.
YOUR BODY IS MADE UP OF CELLS
Trillions of them, in fact.
CHROMOSOMES CONTAIN YOUR UNIQUE DNA
Those helix shapes you see in science books.
YOUR DNA CONTAINS GENES THAT INFLUENCE SPECIFIC TRAITS
From your eye color to your eating habits.

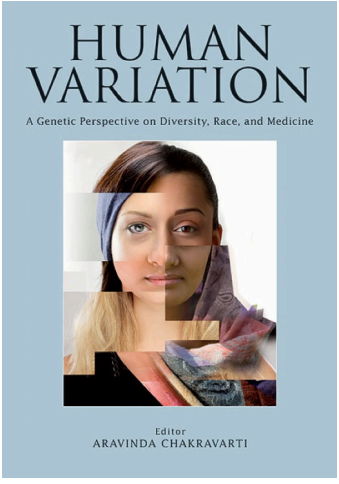
2001 - Publication of the human genome





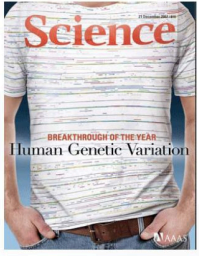
The human genome?

Each individual has genetic differences



2007 SCIENTIFIC BREAKTHROUGH OF THE YEAR

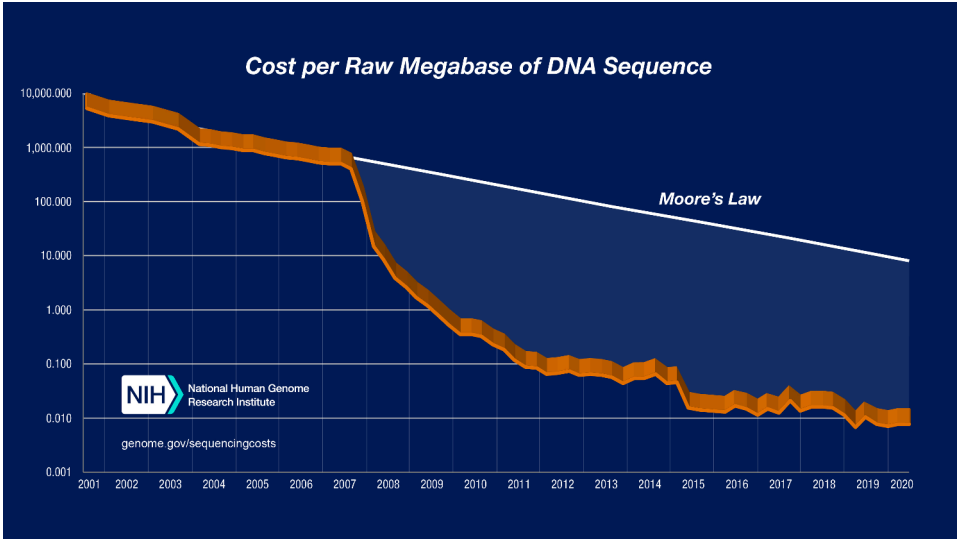
Science Magazine, December 21, 2007



"It's all about me!"

Single Nucleotide Polymorphisms (SNPs)

The cost of DNA sequencing



DNA Sequencing market by application

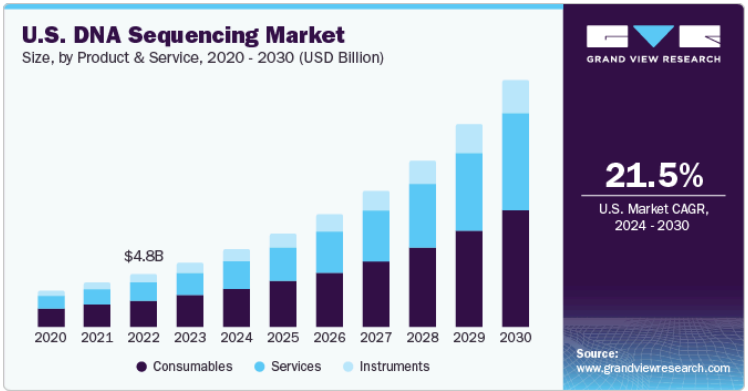
https://www.alliedmarketresearch.com/dna-sequencing-market

The global DNA sequencing market was valued at $4.8 billion in 2022. **Possible application area(s) now and in the future:**

- Diagnostics
- Biomarkers & Cancer
- Reproductive Health
- Personalized Medicine
- Forensics
- Others

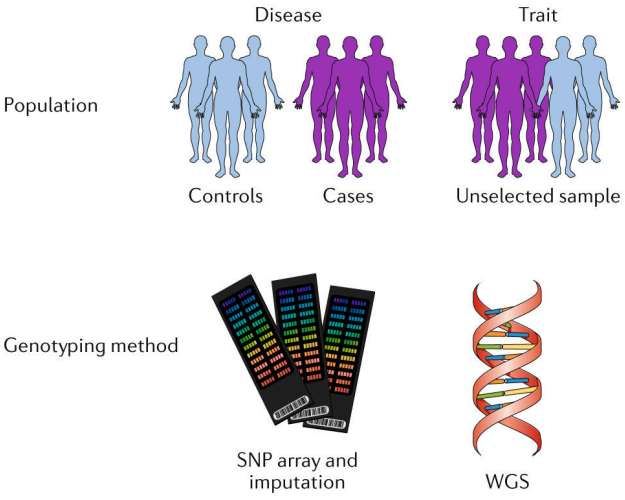
DNA Sequencing market

Substantial growth predicted



Genome-wide association study (GWAS)

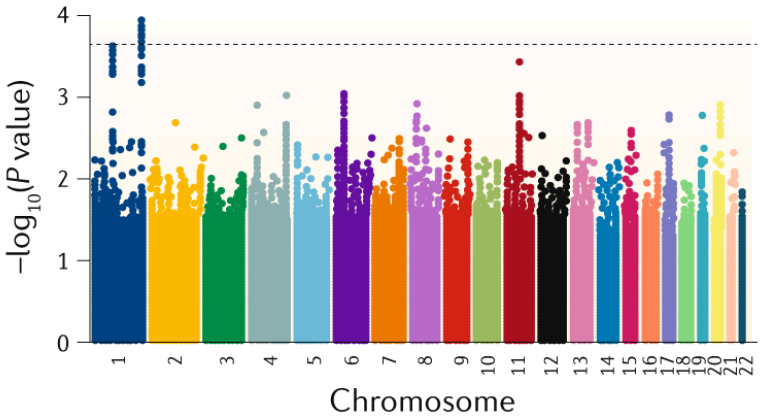
Linking genetic and phenotypic information



Genome-wide association study (GWAS)

Manhatten Plot

Statistical association



Genome-wide association study (GWAS)

Purpose and caveats

GWAS: **Find genetic association with trait (e.g. disease)**

- Most traits are not "simple": Polygenic or environmental component
- Statistical power: Effect size is small or too large
- Functional association of identified SNPs

What Is a Journal Club? (It's not an actual club)

- A journal club is a series of meetings in which somebody is elected to present a research paper, its methods, and findings to a group of colleagues.
- The broad goal to stimulate discussion and ideas that the attendees may apply to their own project.

Journal club

Group discussion of a (recently) published scientific paper

What would be an interesting JC topic to discuss at a course in Singapore?

Journal club

Group discussion of a (recently) published scientific paper

What would be an interesting JC topic to discuss at a course in Singapore?

Genetic basis of income

Hill et al., 2019, Nature Communications

ARTICLE

https://doi.org/10.1038/s41467-019-13585-5 OPEN

Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income

W. David Hill[1,2*], Neil M. Davies[3,4], Stuart J. Ritchie[5], Nathan G. Skene[6,7,8], Julien Bryois[9], Steven Bell[10,11,12], Emanuele Di Angelantonio[10,11,12,13], David J. Roberts[14,15,16], Shen Xueyi[17], Gail Davies[1,2], David C.M. Liewald[1,2], David J. Porteous[1,18], Caroline Hayward[19], Adam S. Butterworth[10,11,12], Andrew M. McIntosh[1,17], Catharine R. Gale[1,2,20,21] & Ian J. Deary[1,2,21]

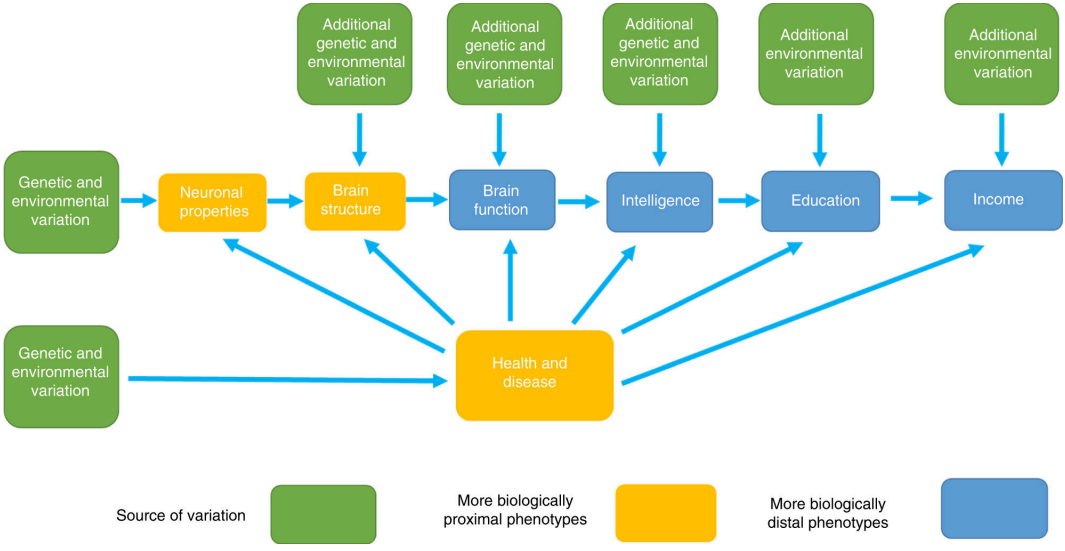
Motivation of the study

What is the genetic basis of socioeconomic position

- **Advantaged socioeconomic backgrounds**
  - live longer
  - have better mental health
  - have better physical health
- The link between socioeconomic position and health is typically thought to be due to **environmental factors**
- **Genetic factors** have been discussed as a partial explanation

Pathway from genetic inheritance to income

Indirect effects of genetic variants on income



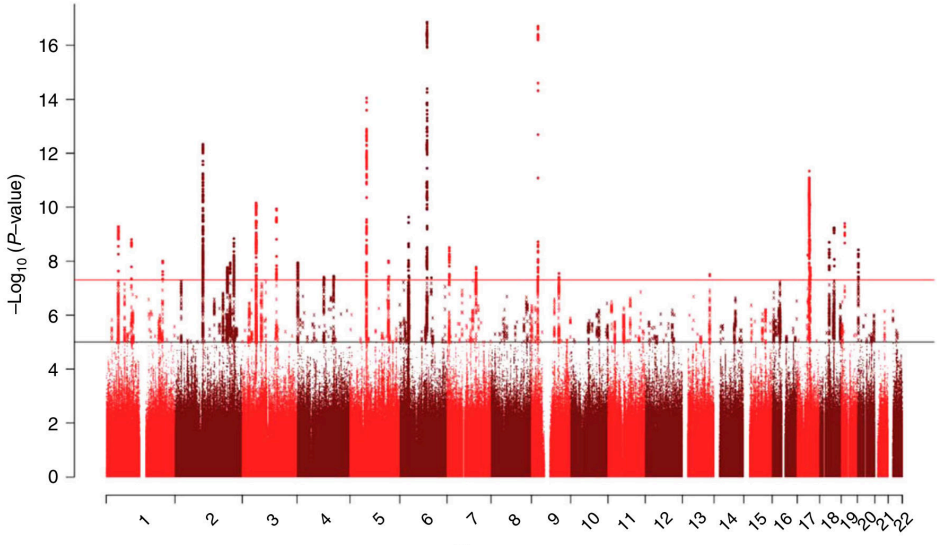
Data in Hill et al.

UK samples were analysed

- **UK BioBank**: an open-access resource established to examine the determinants of disease in middle-aged and older adults living in the United Kingdom
- Recruitment occurred **between 2006 and 2010**
- Analyzed **286,301 participants** aged 39–73 years (51.6% female)
- **5-point scale** of household income (£1 $\approx$ \$1.3)
  1. less than £18,000
  2. £18,000–£29,999
  3. £30,000–£51,999
  4. £52,000–£100,000
  5. greater than £100,000

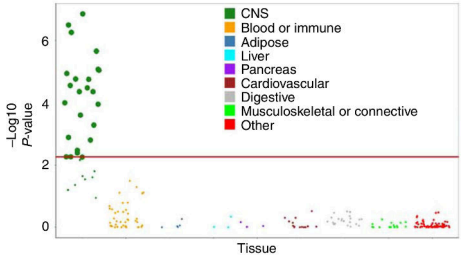
SNP-based analysis of income

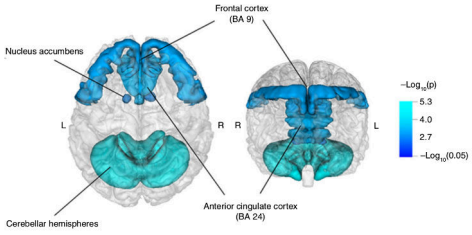
- 3712 SNPs attained genome-wide significance



Potential functional role

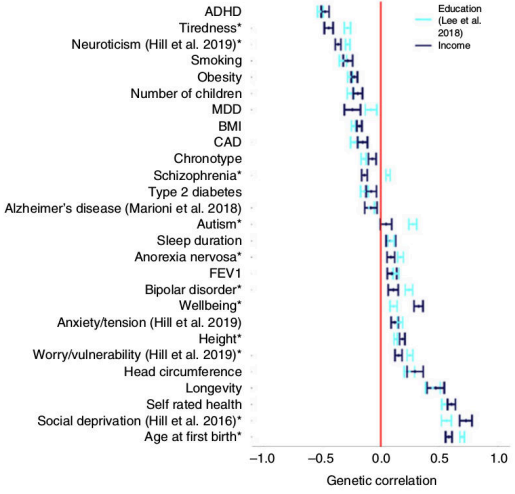
- SNPs and Genes **highly expressed in the brain**





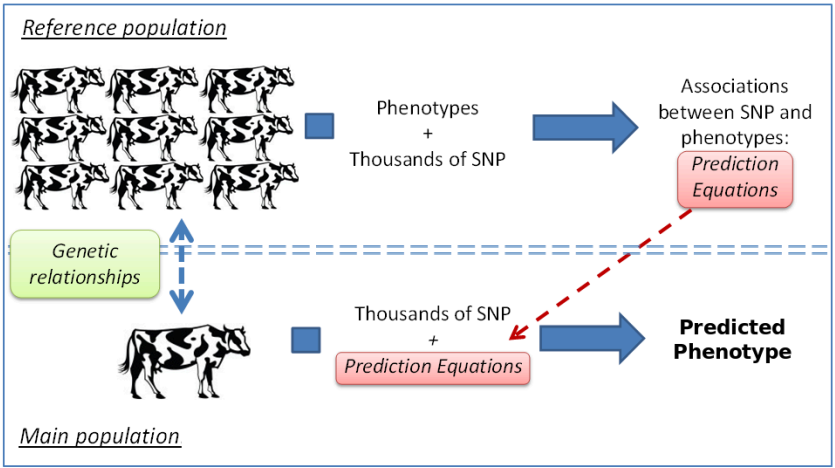
Education and Income versus other traits

Correlation to 27 data sets (psychological traits, mental health, health and well-being, anthropometric traits)



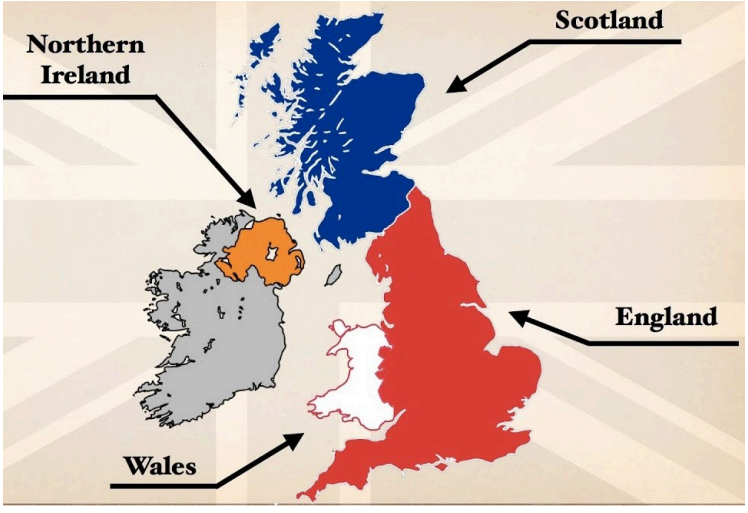
Genomic prediction

A tool used in animal and plant breeding



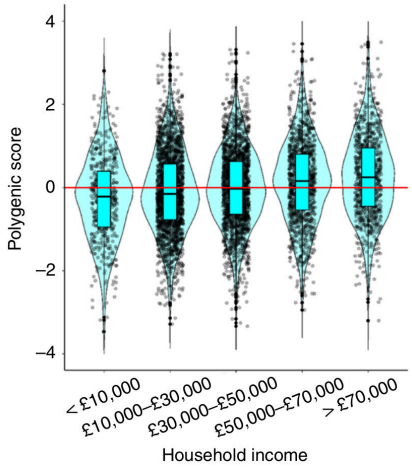
Genomic prediction

UK map



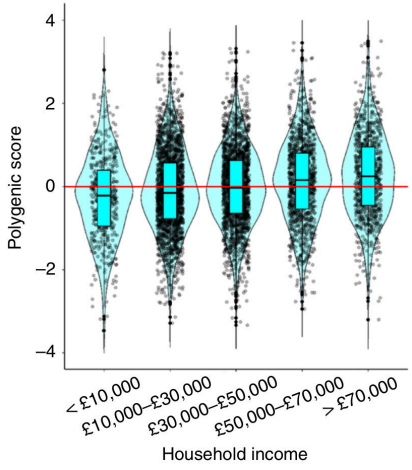
Genomic prediction

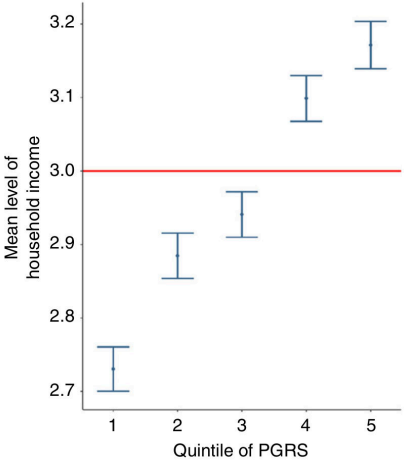
- Polygenic risk scores (PGRSs) were used on Scottish samples



Genomic prediction

- Polygenic risk scores (PGRSs) were used on Scottish samples





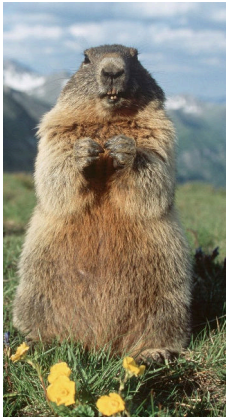
FAQ for Hill et al.

Majority of the reasons people differed in their level of household income were **not genetic**

- Do your results mean that an individual's level of income is determined at birth?
- Have you found "the money gene"?
- Can we now tell what someone's income is from their genes alone?
- What are the practical applications of this research?

Marmota marmota - Alpine marmot genome project [1]

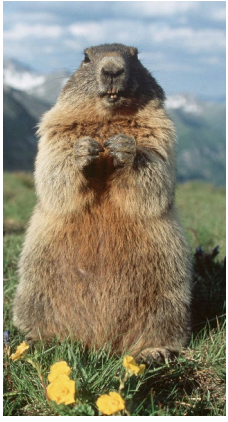
- **Large rodent related to the groundhog**
  - Extremely well adapted to low temperatures
  - Habitant in mountainous areas in central Europe at heights >1,600 metres
  - Up to nine months per year in hibernation
- **Monitored for over 25 years**
  - Study population in France
  - Body mass, litter size and pup winter survival negatively impacted by climate
  - General thinning of snow cover in the last three decades



[1] Gossmann et al., 2019, *Current Biology*; Gossmann and Ralser, 2020, *Trends in Genetics*

Marmota marmota - Alpine marmot genome project [1]

- **Large rodent related to the groundhog**
  - Extremely well adapted to low temperatures
  - Habitant in mountainous areas in central Europe at heights >1,600 metres
  - Up to nine months per year in hibernation
- **Monitored for over 25 years**
  - Study population in France
  - Body mass, litter size and pup winter survival negatively impacted by climate
  - General thinning of snow cover in the last three decades

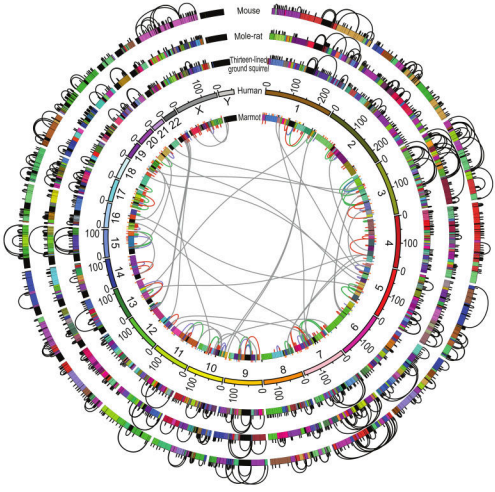


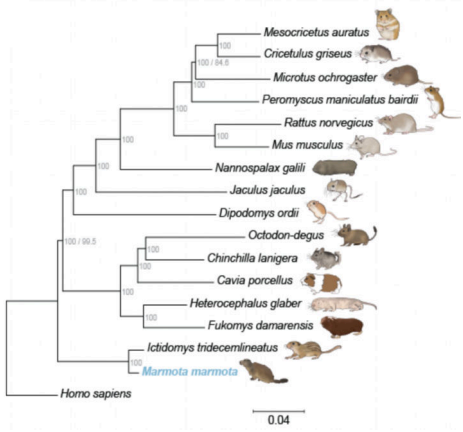
Ideal species to study adaptation to extreme climates

[1] Gossmann et al., 2019, *Current Biology*; Gossmann and Ralser, 2020, *Trends in Genetics*

Comparative genomic features

Chromosome synteny and phylogenetic relationship

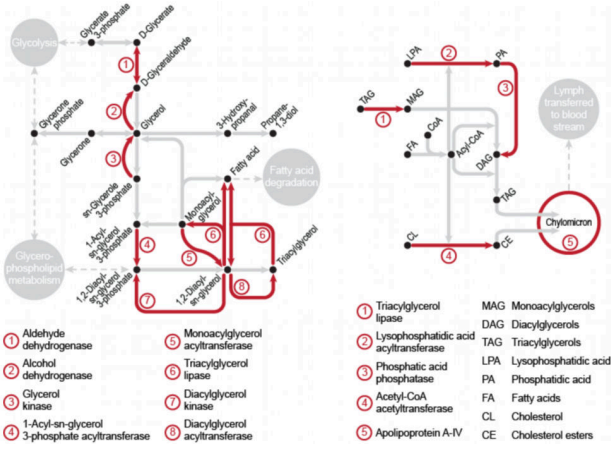




What genes show signatures of rapid evolution?

Genes involved in fat synthesis and storage

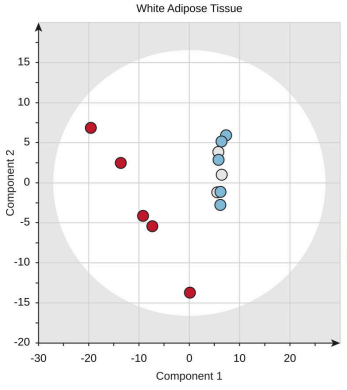
- Diglyceride (DAG) and triglyceride (TAG) biosynthesis
- Fat storage



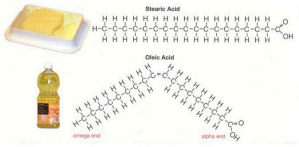
Lipidomics: Adipose tissue lipid composition

Mass spectrometry and liquid chromatography

Enrichment of unsaturated FAs with longer chains and more double bounds in marmot adipose tissue compared to mouse and rat

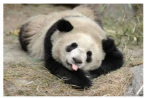


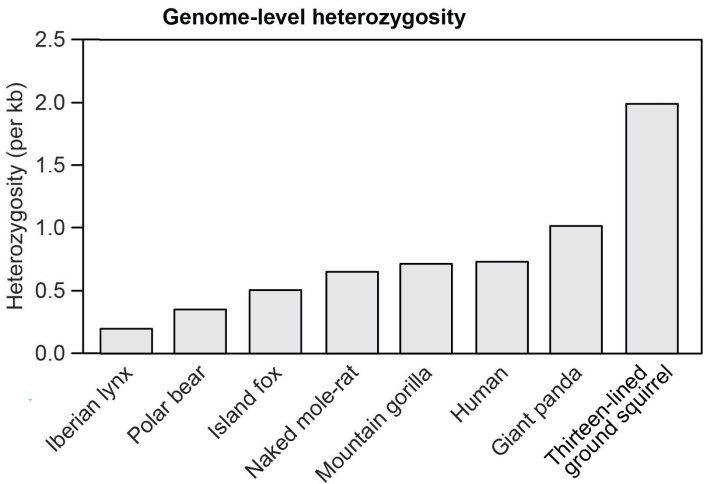
	Marmot	Mouse	Rat
Average chain length of TAGs	53.6	51.0	49.7
Average double bounds in TAGs	6.6	2.8	2.7
Average chain length in DAGs	44.0	39.0	37.5
Average double bonds in DAGs	8.0	2.0	2.5



Genomic diversity of (endangered) species

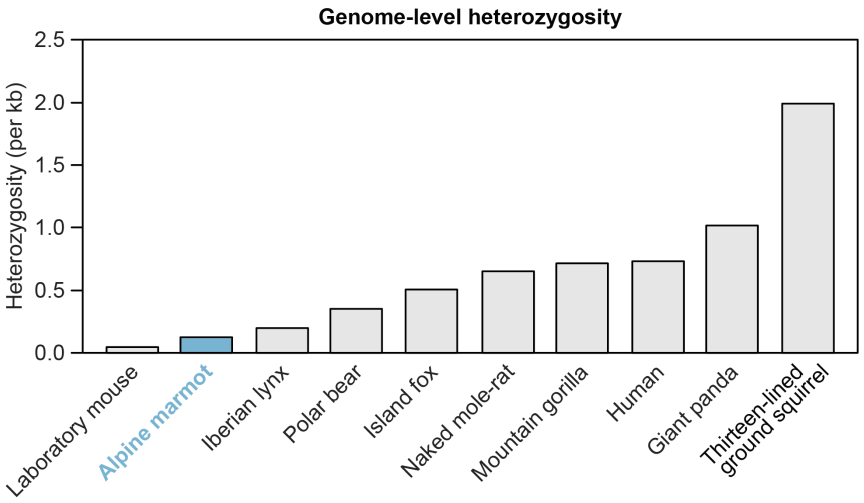
Genome-wide heterozygosity is very low for some highly threatened species





Extreme low diversity of the marmot reference genome

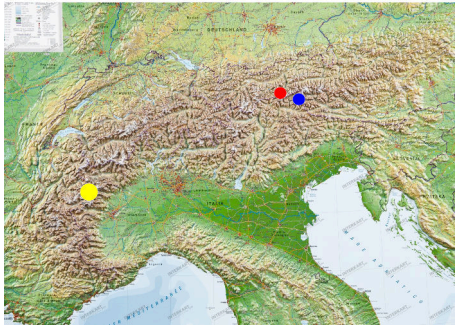
Genome-wide heterozygosity is even lower than for highly threatened species



Population genetic analysis

Whole genome re-sequencing from 2 further populations

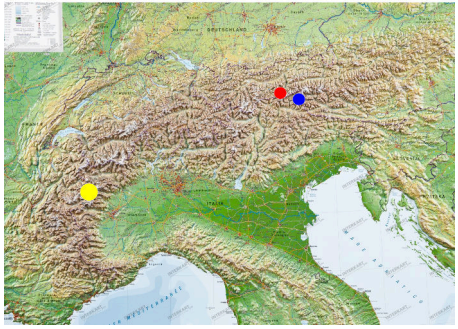
- **8 additional individuals**
  - Reference genome from **Mauls**/Italy
  - 4 from **Gsies**/Italy
  - 4 from France (LGS)



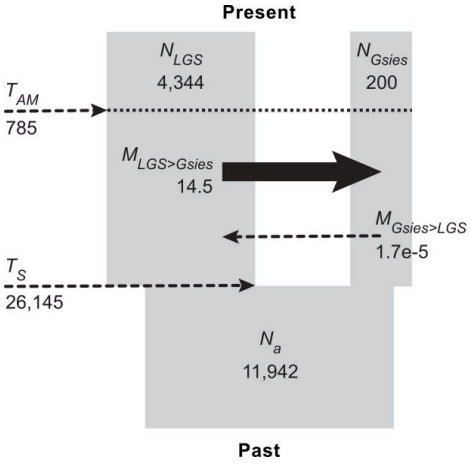
Population genetic analysis

Whole genome re-sequencing from 2 further populations

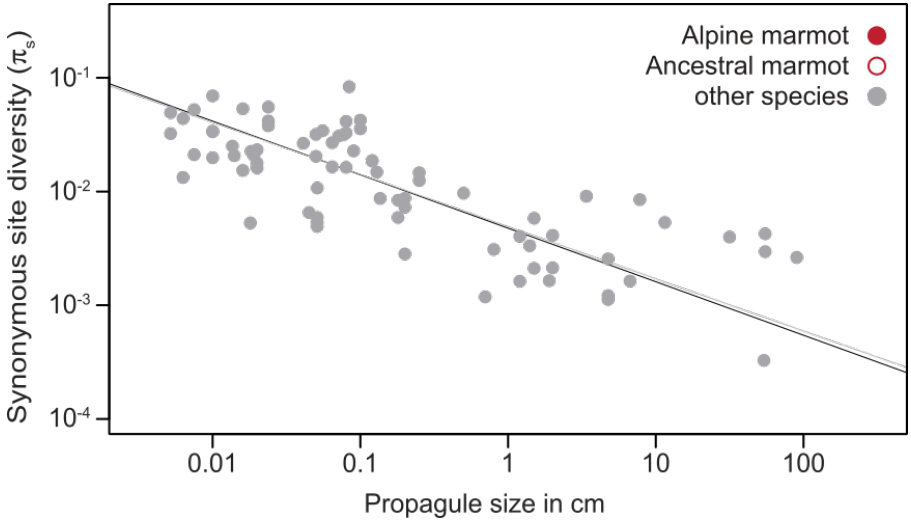
- **8 additional individuals**
  - Reference genome from **Mauls**/Italy
  - 4 from **Gsies**/Italy
  - 4 from France (LGS)



- **Demographic analysis**

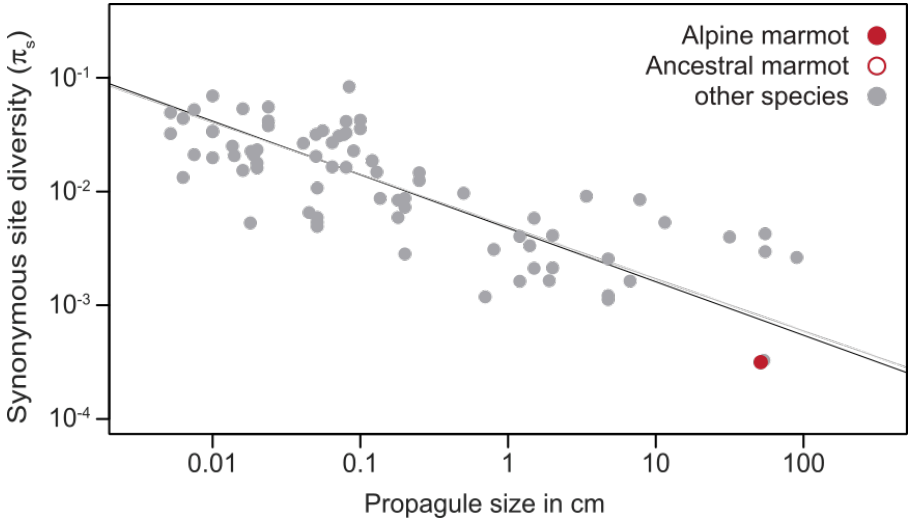


Propagule size is a very good predictor of genetic diversity[1]



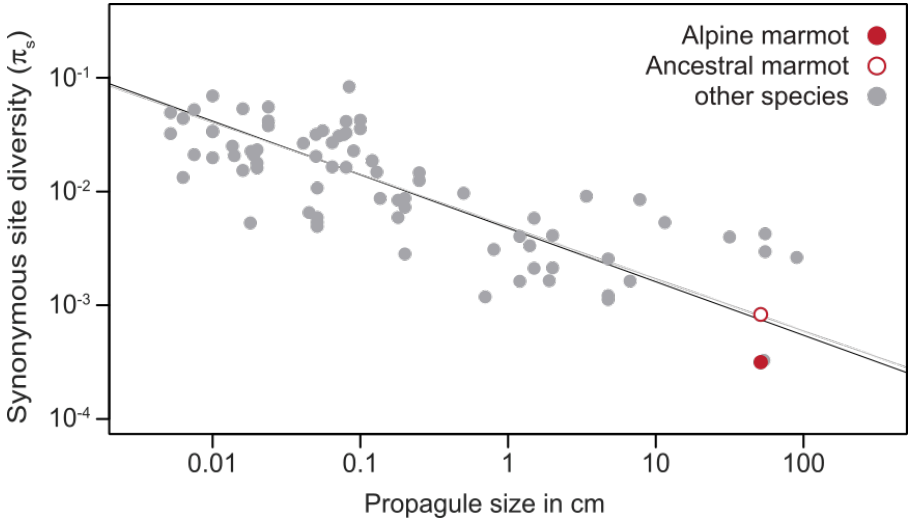
[1] Romiguier et al., 2014

Propagule size is a very good predictor of genetic diversity[1]



[1] Romiguier et al., 2014

Propagule size is a very good predictor of genetic diversity[1]



[1] Romiguier et al., 2014

Summary

- Despite being highly abundant and well adapted, the Alpine marmot is **among the least genetically diverse animals**
- Low genetic diversity is a consequence of consecutive, **climate-related events** (long-term extreme niche adaptation, and a recent bottleneck after the last ice age)
- **Large population** size can coexist with very low genetic variation

Future: Why bother about genomic and other omic research?

- Genomic and other type of omics data will become increasingly available
  - Data amount of unprecedented scale
  - Novel data (e.g. Pac-Bio, Nanopore)
- Possibility to generate "own" pinpointed data
- Need for **novel** and **innovative** approaches
  - Technical
  - **Biological-driven research questions**

What have you learned today

- BCI @ TU Dortmund: An exciting and diverse place
- Importance of OMICs
- How a GWAS works
- The special genome of the Alpine marmot

What have you learned today

- BCI @ TU Dortmund: An exciting and diverse place
- Importance of OMICs
- How a GWAS works
- The special genome of the Alpine marmot