

Zmienność genetyczna człowieka

Cechy wieloczynnikowe - od asocjacji do wielogenowych wskaźników ryzyka

Różnorodność genetyczna człowieka

- Projekt “1000 genomów” – poszukiwanie różnic w genomach różnych ludzi (2500 osób)
- Dane (2015):
 - 84,7 mln. miejsc zmiennych (SNP)
 - 3,6 mln. indeli
 - 60 000 wariantów strukturalnych (duże delecje/insercje, rearanżacje)

A global reference for human genetic variation

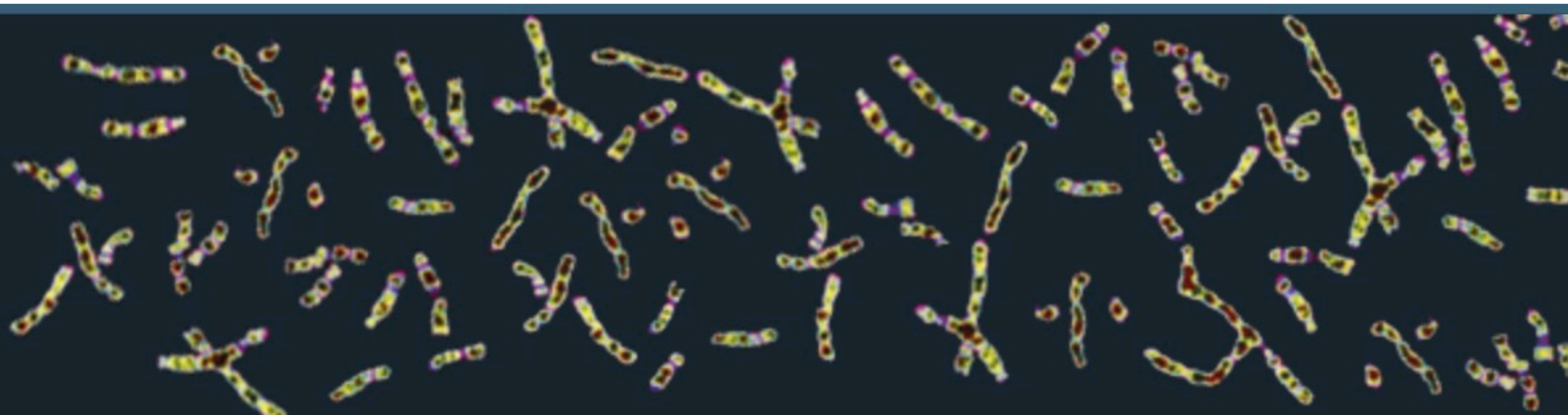
The 1000 Genomes Project Consortium*

NATURE | VOL 526 | 1 OCTOBER 2015

- Więcej nukleotydów, niż w całym genomie drożdży

IGSR: The International Genome Sample Resource

Providing ongoing support for the 1000 Genomes Project data



[Home](#) [About](#) [Data](#) [Portal](#) [Analysis](#) [Contact](#) [Browser](#) [FAQ](#)

Q

IGSR and the 1000 Genomes Project



Populations: ● - African; ● - American; ● - East Asian; ● - European; ● - South Asian;

The International Genome Sample Resource (IGSR) was established to ensure the ongoing usability of data generated by the 1000 Genomes Project and to extend the data set. More information is available [about the IGSR](#).

Links

- [Announcements](#)
- [IGSR Sample Collection Principles](#)
- [1000 Genomes Project Publications](#)

- [File formats](#)
- [Software tools](#)
- [Download data](#)
- [User survey](#)
- [Twitter](#)

<http://www.internationalgenome.org/>

Table 1 | Median autosomal variant sites per genome

	AFR		AMR		EAS		EUR		SAS	
Samples	661		347		504		503		489	
Mean coverage	8.2		7.6		7.7		7.4		8.0	
	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons	Var. sites	Singletons
SNPs	4.31M	14.5k	3.64M	12.0k	3.55M	14.8k	3.53M	11.4k	3.60M	14.4k
Indels	625k	-	557k	-	546k	-	546k	-	556k	-
Large deletions	1.1k	5	949	5	940	7	939	5	947	5
CNVs	170	1	153	1	158	1	157	1	165	1
MEI (Alu)	1.03k	0	845	0	899	1	919	0	889	0
MEI (L1)	138	0	118	0	130	0	123	0	123	0
MEI (SVA)	52	0	44	0	56	0	53	0	44	0
MEI (MT)	5	0	5	0	4	0	4	0	4	0
Inversions	12	0	9	0	10	0	9	0	11	0
Nonsynon	12.2k	139	10.4k	121	10.2k	144	10.2k	116	10.3k	144
Synon	13.8k	78	11.4k	67	11.2k	79	11.2k	59	11.4k	78
Intron	2.06M	7.33k	1.72M	6.12k	1.68M	7.39k	1.68M	5.68k	1.72M	7.20k
UTR	37.2k	168	30.8k	136	30.0k	169	30.0k	129	30.7k	168
Promoter	102k	430	84.3k	332	81.6k	425	82.2k	336	84.0k	430
Insulator	70.9k	248	59.0k	199	57.7k	252	57.7k	189	59.1k	243
Enhancer	354k	1.32k	295k	1.05k	289k	1.34k	288k	1.02k	295k	1.31k
TFBSs	927	4	759	3	748	4	749	3	765	3
Filtered LoF	182	4	152	3	153	4	149	3	151	3
HGMD-DM	20	0	18	0	16	1	18	2	16	0
GWAS	2.00k	0	2.07k	0	1.99k	0	2.08k	0	2.06k	0
ClinVar	28	0	30	1	24	0	29	1	27	1

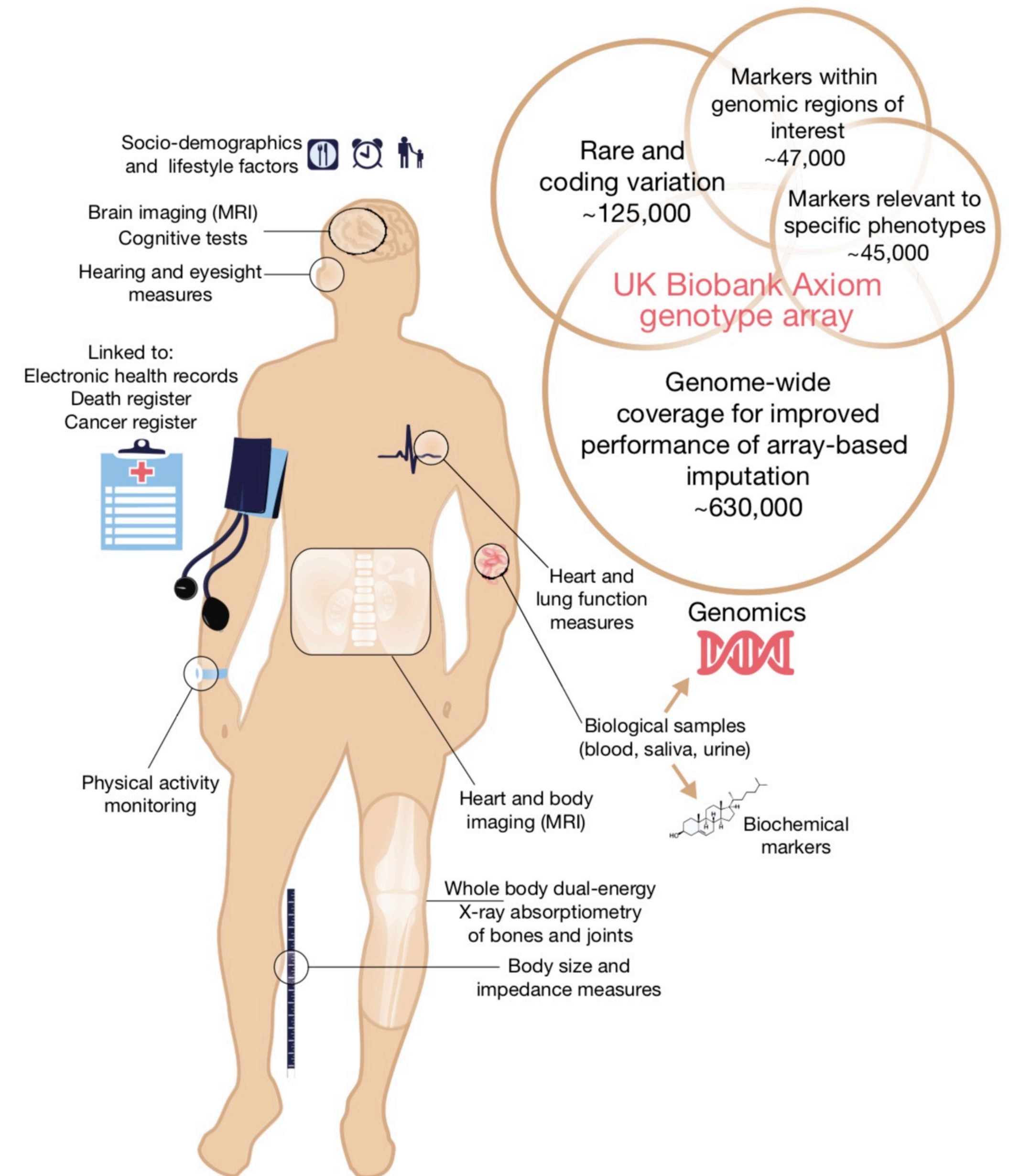
Nowe dane

- Obecnie koszt sekwencjonowania pojedynczego genomu <1000\$
- UK BioBank: ~500 000 osób

The UK Biobank resource with deep phenotyping and genomic data

Clare Bycroft^{1,13}, Colin Freeman^{1,13}, Desislava Petkova^{1,12,13}, Gavin Band¹, Lloyd T. Elliott², Kevin Sharp², Allan Motyer³, Damjan Vukcevic^{3,4}, Olivier Delaneau^{5,6,7}, Jared O'Connell⁸, Adrian Cortes^{1,9}, Samantha Welsh¹⁰, Alan Young¹¹, Mark Effingham¹⁰, Gil McVean^{1,11}, Stephen Leslie^{3,4}, Naomi Allen¹¹, Peter Donnelly^{1,2,14} & Jonathan Marchini^{1,2,14*}

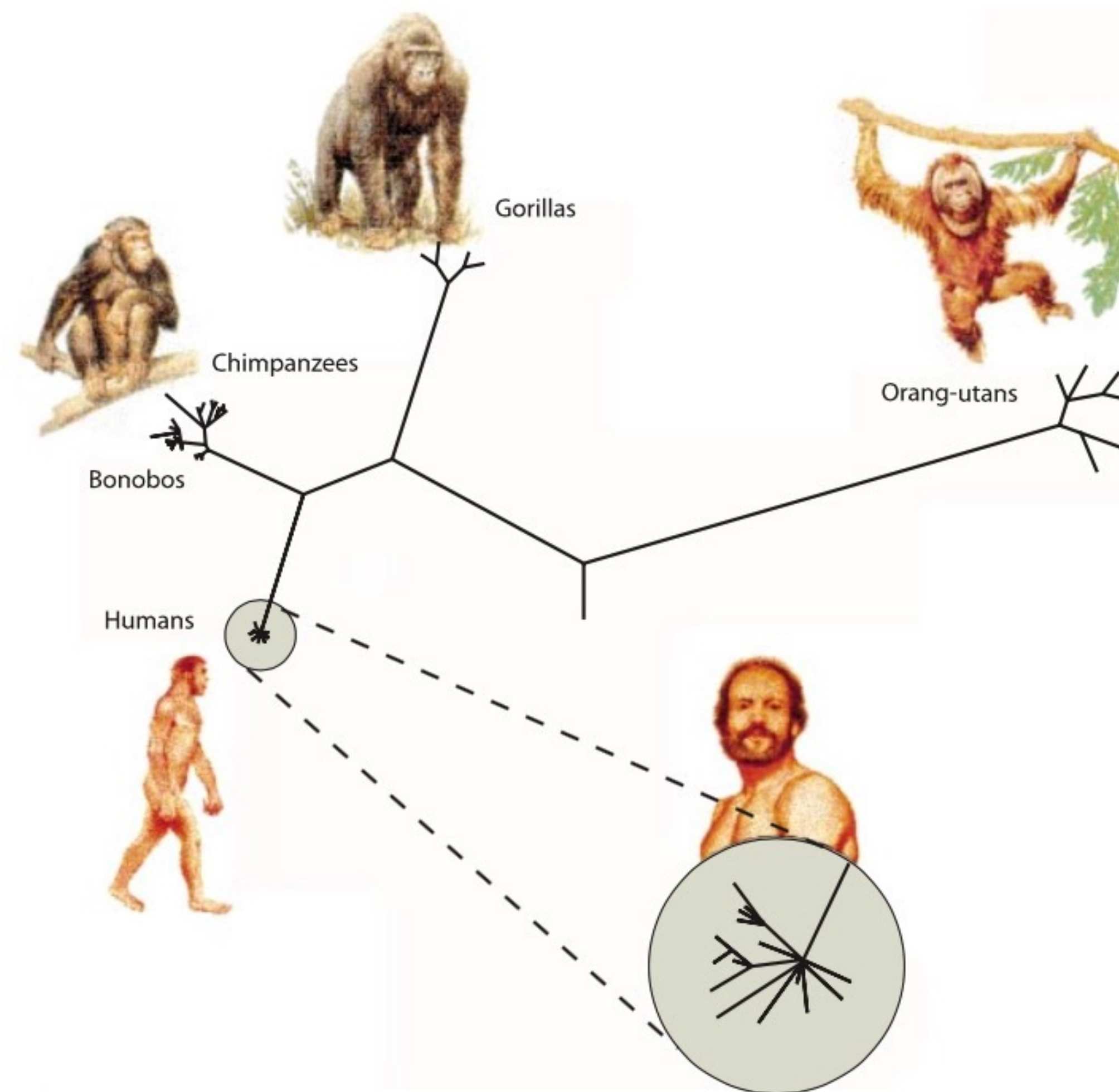
11 OCTOBER 2018 | VOL 562 | NATURE | 207



Metody badania

- Mikromacierze - panele znanych SNP
 - zaleta: standaryzacja
- Sekwencjonowanie
 - zaleta: wykrywanie alleli bardzo rzadkich

Różnorodność genetyczna ludzi jest stosunkowo niewielka



Różnorodność genetyczna ludzi jest stosunkowo niewielka

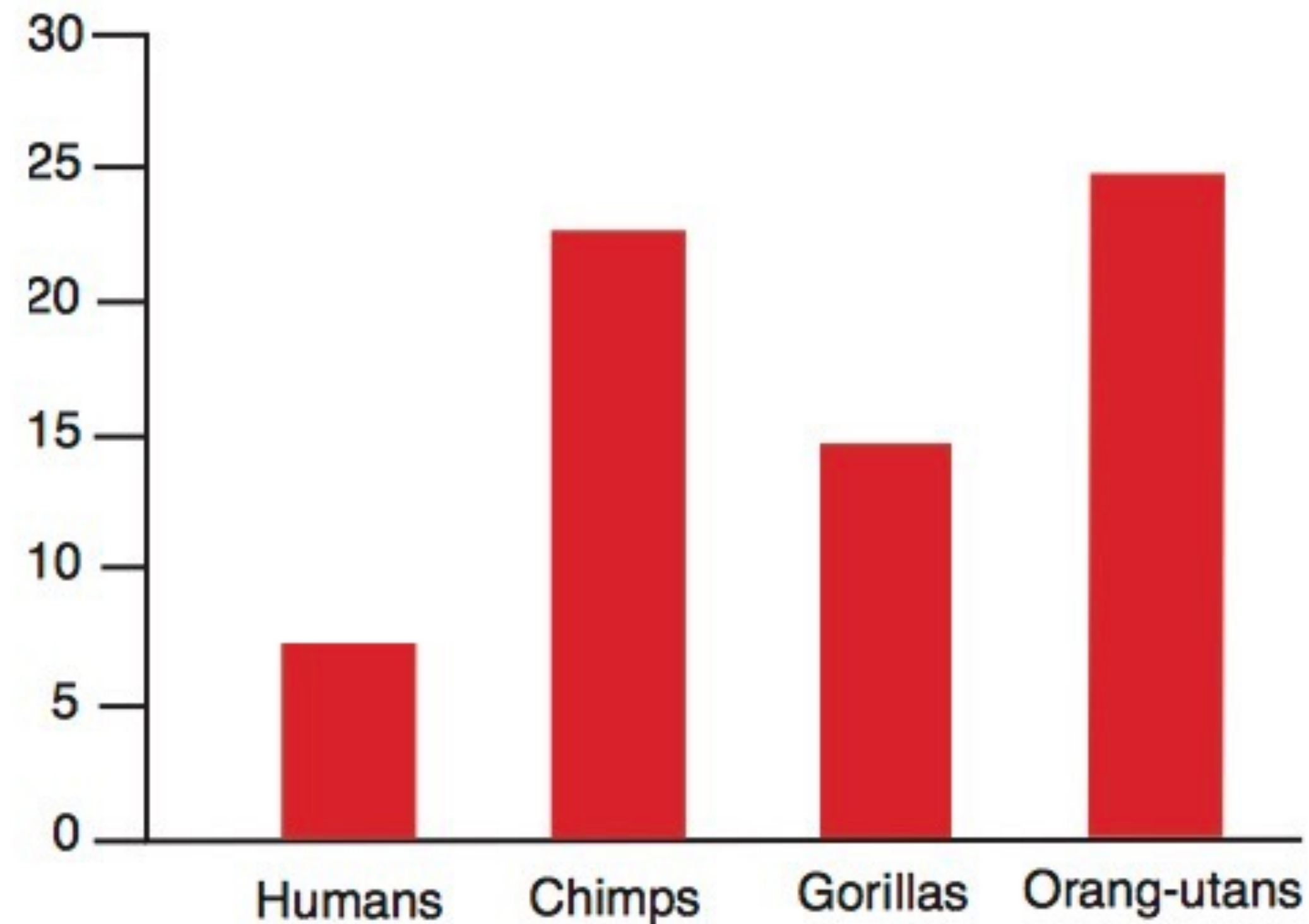


Fig. 11 DNA sequence diversity within humans and great apes. Values are based on the number of variable positions within each species taking the number of sequences determined into account (Watterson's diversity estimator, θ_w).

Jak dziedziczą się cechy człowieka?

- Mendlowskie - jeden gen = jedna cecha - **jednogenowe**
 - zmienność konkretnego genu decyduje o fenotypie konkretnej cechy
 - w przypadku chorób - tzw. mutacje sprawcze
- **Oligogenowe** - zależne od współdziałania kilku - kilkunastu genów
 - np. kolor oczu, włosów
- **Wieloczynnikowe** - zależą od współdziałania wielu (setek, tysięcy) genów i środowiska

Jak dziedziczą się cechy człowieka?

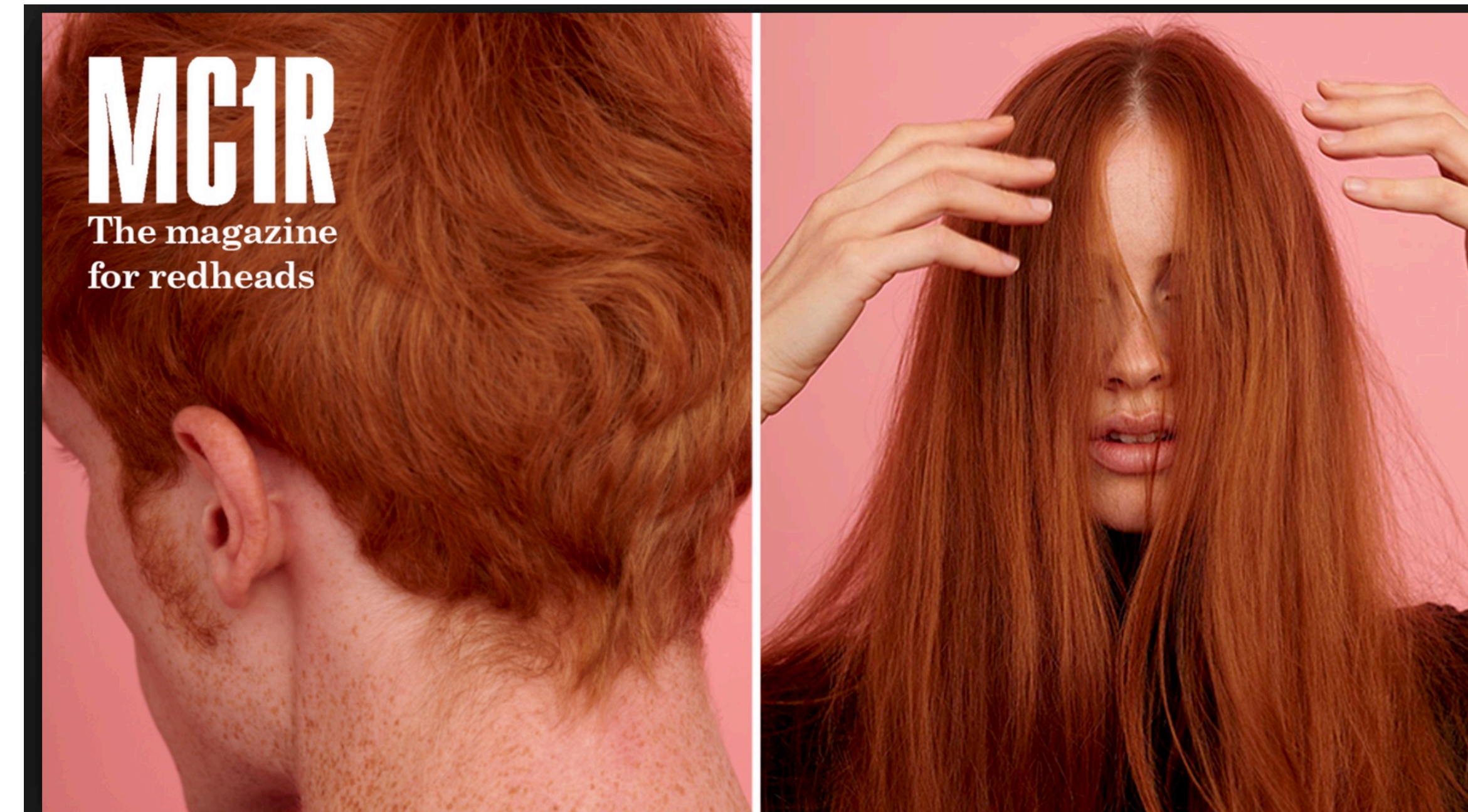
- Mendlowskie - jeden gen = jedna cecha
 - dobrze potrafimy identyfikować geny, analizować dziedziczenie, wykrywać mutacje i przewidywać fenotyp, możliwe modyfikacje
- Oligogenowe - zależne od działania kilku - kilkunastu genów
 - potrafimy analizować i przewidywać fenotyp, ale nie ze 100% dokładnością
- Wieloczynnikowe - zależą od współdziałania wielu (setek, tysięcy) genów i środowiska
 - nie potrafimy dobrze przewidywać, dopiero zaczynamy poznawać złożoność i odkrywać korelacje statystyczne

Jak dziedziczą się cechy człowieka?

- Mendlowskie - jeden gen = jedna cecha
 - znamy wiele chorób, które tak się dziedziczą, ale są to choroby rzadkie
 - tylko pojedyncze przykłady cech zmienności prawidłowej
- Wieloczynnikowe - zależą od współdziałania wielu (setek, tysięcy) genów i środowiska
 - praktycznie wszystkie aspekty zmienności prawidłowej
 - większość często występujących chorób

Cechy jednogenowe zmienności prawidłowej

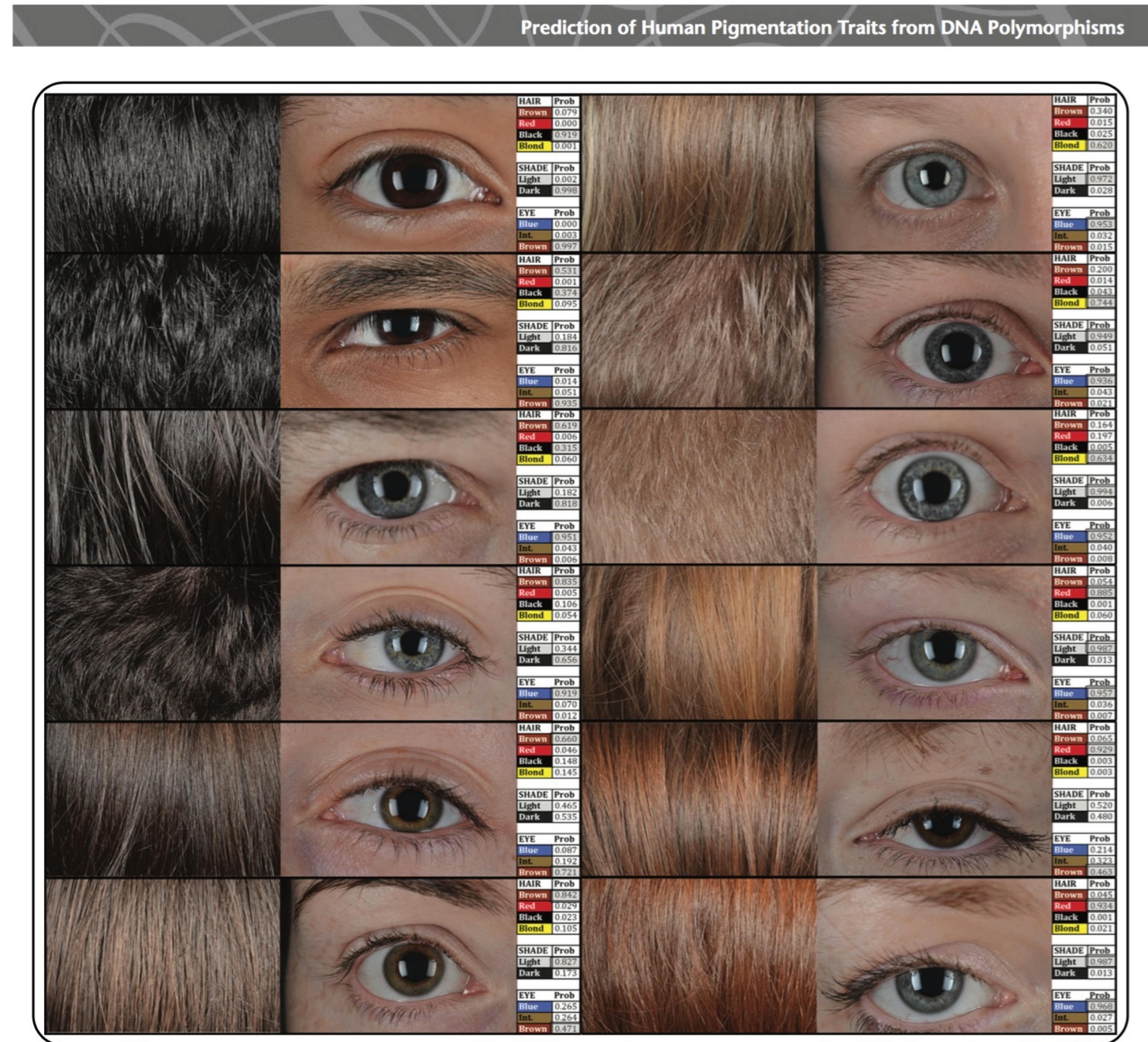
- Bardzo nieliczne przykłady
- Większość “podręcznikowych” przykładów to uproszczenia (np. płatek ucha, zwijanie języka)
- Przykłady:
 - M1CR - receptor melanokortyny 1 (receptor MSH), rude włosy, jasna skóra, piegi, podatność na czerniaka
 - OR6A2 - wariant genu (~17% w Europie) powoduje negatywną reakcję smakową na związki (aldehydy) występujące w świeżej naci kolendry



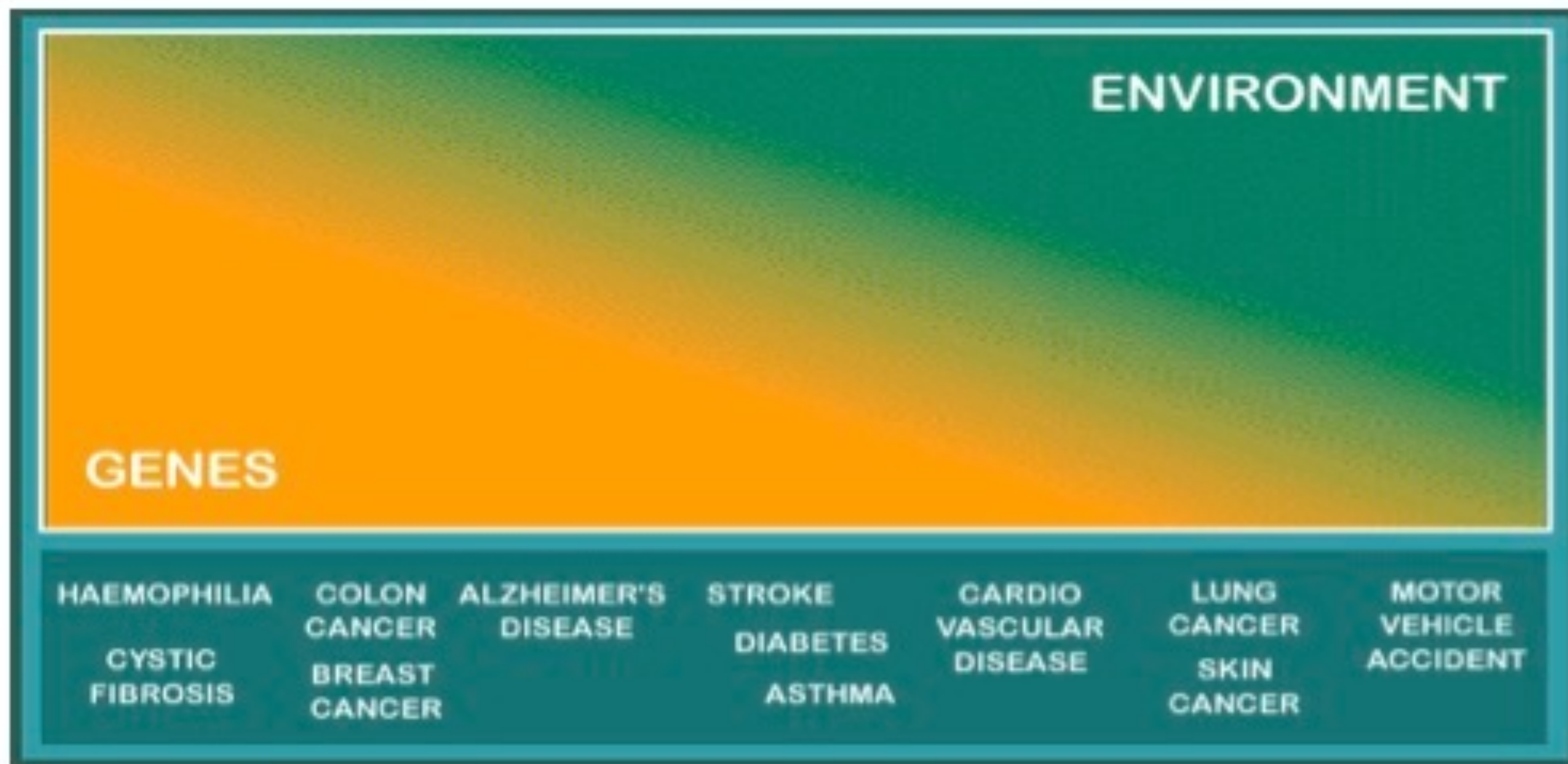
Cechy oligogenowe - pigmentacja

- System HirisPlex
- 24 loci
- przewidywanie koloru oczu, włosów i pigmentacji skóry
- >86% dokładność
 - najwyższa dla włosów rudych, najniższa dla jasnych
- Genetyczny rysopis
 - można też przewidywać:
 - skłonność do łysienia
 - wiek (na pods. metylacji DNA)

Branicki & Keyser, 2015



Cechy wieloczynnikowe



Na przykład wzrost

- W determinowaniu wzrostu człowieka bierze udział około 200 genów



Defekty a zmienność prawidłowa

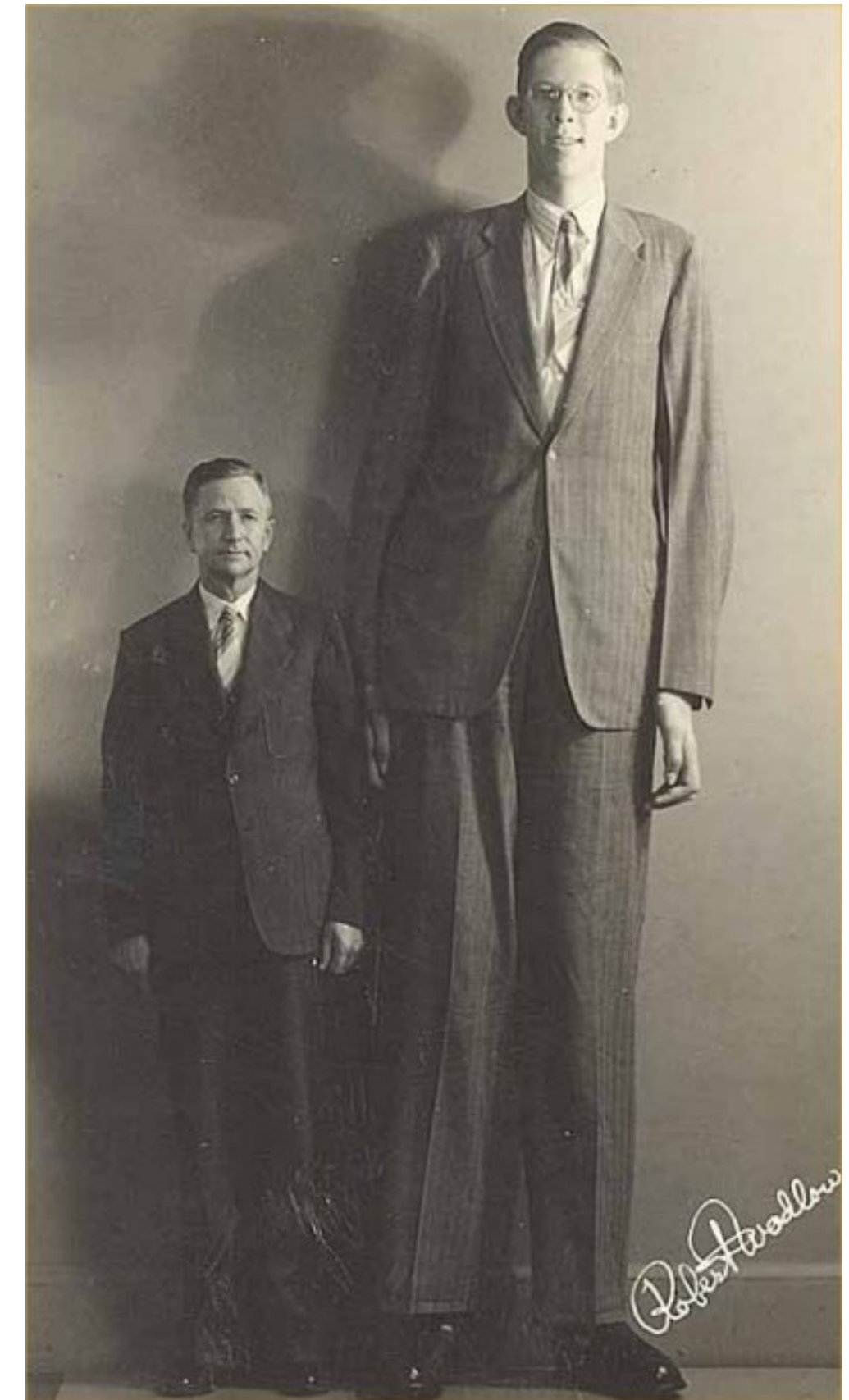
- Achondroplazja
- Mutacja pojedynczego genu *FGFR3* u człowieka
- cecha mendłowska, dominująca (letalna u homozygot)
- niski wzrost, nieproporcjonalnie krótkie kończyny



Defekty a zmiennność prawidłowa

- Czy gen *FGFR3* wystarczy do zrozumienia genetyki różnic wzrostu u ludzi?
- Nie - to nie jest cecha jednogenowa
- W odziedziczną zmiennność wzrostu zaangażowane jest ponad 200 genów

GIANT (Genetic Investigation of Anthropometric Traits),
Lango et al. Nature. 2010 467(7317):832-8.



Defekty a zmienność prawidłowa

- To, że mutacja jakiegoś pojedynczego genu całkowicie zaburza działanie jakiegoś systemu nie oznacza, że
 - jest on jedynym czynnikiem odpowiadającym za ten system
 - odgrywa on istotną rolę w prawidłowej (populacyjnej) zmienności tej cechy

Cechy wieloczynnikowe

- Częste choroby, odziedziczalna część ryzyka
 - np. nowotwory (nie dziedziczne)
 - choroby serca i układu naczyniowego
 - cukrzyca
 - choroby psychiczne
- Prawidłowa zmienność fenotypowa
- Za odziedziczalność muszą też odpowiadać częste warianty sekwencyjne (częste polimorfizmy)

Cechy wieloczynnikowe a mendlowskie

Mendlowskie	Wieloczynnikowe
Proste dziedziczenie, allele pojedynczych genów, wysoka penetracja	Złożone dziedziczenie: wiele czynników (genetycznych i środowiskowych), każdy z nich osobno ma niską penetrację
Mutacja <u>powoduje</u> chorobę	Polimorfizmy genów wpływają na ryzyko (zwiększają albo zmniejszają)
Badanie: analiza sprzężeń, sekwencjonowanie	Badanie: metody statystyczne, głównie nieparametryczne
Rzadkie choroby	Wiele częstych chorób i cech prawidłowych

Allele - terminologia

- Częste (*common*): $>5\%$ częstość allelu rzadszego (MAF - *Minor Allele Frequency*)
 - praktycznie nie ma wśród nich alleli sprawczych
- Niskiej częstości (*low-frequency*): $5\% > \text{MAF} > 0,5\%$
- Rzadkie (*rare*): $\text{MAF} < 0,5\%$

Genetyka częstych chorób

- Za dziedziczną składową częstych chorób odpowiadają głównie polimorfizmy z $MAF > 0,5\%$, ale efekty każdego z nich pojedynczo są niewielkie
- Rzadkie ($MAF < 0,5\%$) allele o większym ryzyku
 - Rzadkie allele o większym ryzyku odpowiadają za 1-10% przypadków (zależnie od choroby) i mogą dawać przypadki rodzinne
 - np. rodzinne: hipercholesterolemia, nowotwory, Alzheimer

Odziedziczalność

- Za każdy fenotyp odpowiada interakcja genotypu ze środowiskiem
- Odziedziczalność (H^2): proporcja zmienności fenotypowej wyjaśnianej zmiennością genetyczną w populacji
- Niekiedy wyróżniana odziedziczalność *sensu stricto* (h^2): rola addytywnych czynników genetycznych
- Odziedziczalność GWAS: wyjaśniana przez polimorfizmy nukleotydowe w populacji

Różnice między odziedziczalnością a ryzykiem indywidualnym

- Mutacje *BRCA1* i *BRCA2* odpowiadają za ~5% przypadków raka piersi
 - odziedziczalność w populacji jest nieznaczna
- Osoby z mutacjami: istotny wzrost ryzyka zachorowania

Odziedziczalność - oszacowanie

- badania bliźniąt
 - monozygotyczne (MZ) vs. dizygotyczne (DZ)
- agregacja rodzinna
 - częstość objawów u krewnych I stopnia przewyższa obserwowaną u dalszych krewnych i osób niespokrewnionych

Nieporozumienia dotyczące odziedziczalności

- Nie jest to miara udziału cech dziedzicznych w fenotypie
 - jest to miara udziału cech dziedzicznych w zmienności fenotypu
 - kowariancja genotypu i fenotypu

Nieporozumienia dotyczące odziedziczalności

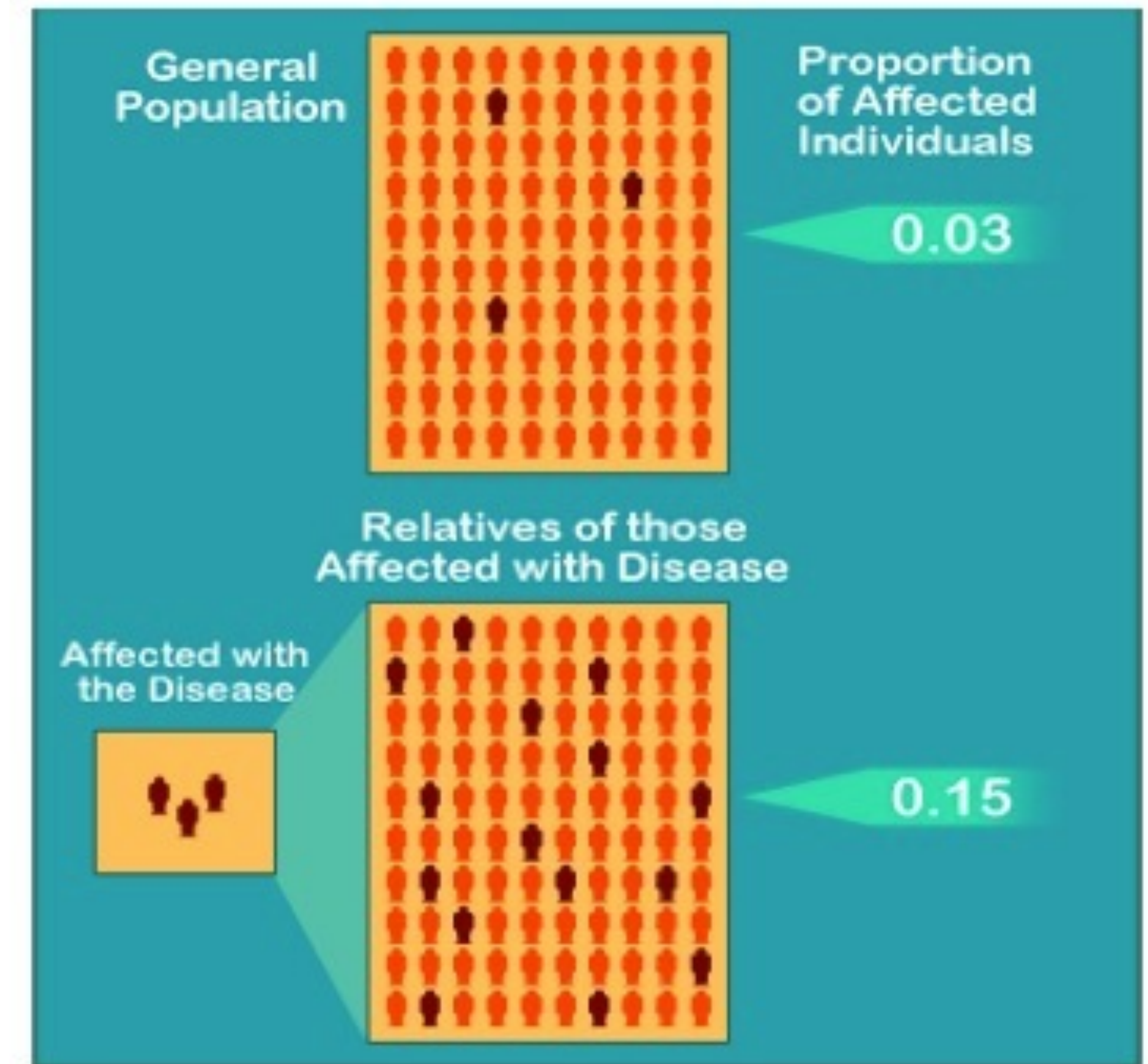
- Nie jest to miara podobieństwa potomstwa i rodziców
 - np. typowe choroby autosomalne recesywne: chore dzieci zdrowych rodziców, 100% odziedziczalność

Nieporozumienia dotyczące odziedziczalności

- Nie jest to własność cechy ani osobnika, tylko populacji
 - zależy od struktury populacji, środowiska itp.
- W populacji jednolitej genetycznie (klony) całość zmienności cechy jest spowodowana zmianami środowiska - 0% odziedziczalności
- wzrost: odziedziczalność zależy od dobrobytu populacji (wysoka w populacjach dobrze odżywionych, niska w czasach głodu)

Agregacja rodzinna

- Czy cecha występuje częściej u krewnych osób mających tę cechę



Agregacja rodzinna

- Choroby afektywne (ostre cykliczne zaburzenia nastroju)
 - jednobiegunowa (epizody depresyjne)
 - dwubiegunowa (epizody maniakalne i depresyjne)

Agregacja rodzinna

- Choroba jednobiegunowa
 - ryzyko w populacji ~3 % (mężczyźni) ~5-9% (kobiety)
 - krewni I stopnia osoby chorej – ryzyko ~10%
 - stosunkowo mała odziedziczalność
- Choroba dwubiegunowa
 - ryzyko w populacji ~1%
 - krewni I stopnia osoby chorej – ryzyko ~ 20%
 - istotna odziedziczalność

Agregacja rodzinna

$$\lambda_S = \frac{\text{ryzyko dla krewnych 1 stopnia chorych}}{\text{ryzyko w populacji}}$$

λ_S – *recurrence risk ratio*

Np.

populacja	0,8
krewni I stopnia chorych	8,5

$$\lambda_S = 10,6$$

Agregacja rodzinna

- Przykładowe wartości λ_s
- Choroby mendlowskie:
 - mukowiscydoza: $0,25/0,0004 = 500$
 - choroba Huntingtona: $0,5/0,0001 = 5000$
- Choroby wieloczynnikowe
 - autyzm: ~ 110
 - stwardnienie rozsiane: ~ 25
 - schizofrenia: ~ 10
 - cukrzyca typ I (IDDM): ~ 15
 - cukrzyca typ II (NIDDM): $\sim 3,5$

Asocjacja

- Nieprzypadkowe współwystępowanie czynników (alleli i fenotypów) na poziomie populacji
 - Czy zawsze asocjacja oznacza zależność przyczynową?
 - Czy każda asocjacja ma wartość diagnostyczną?
 - Czy asocjacja odkrywa “gen na”?

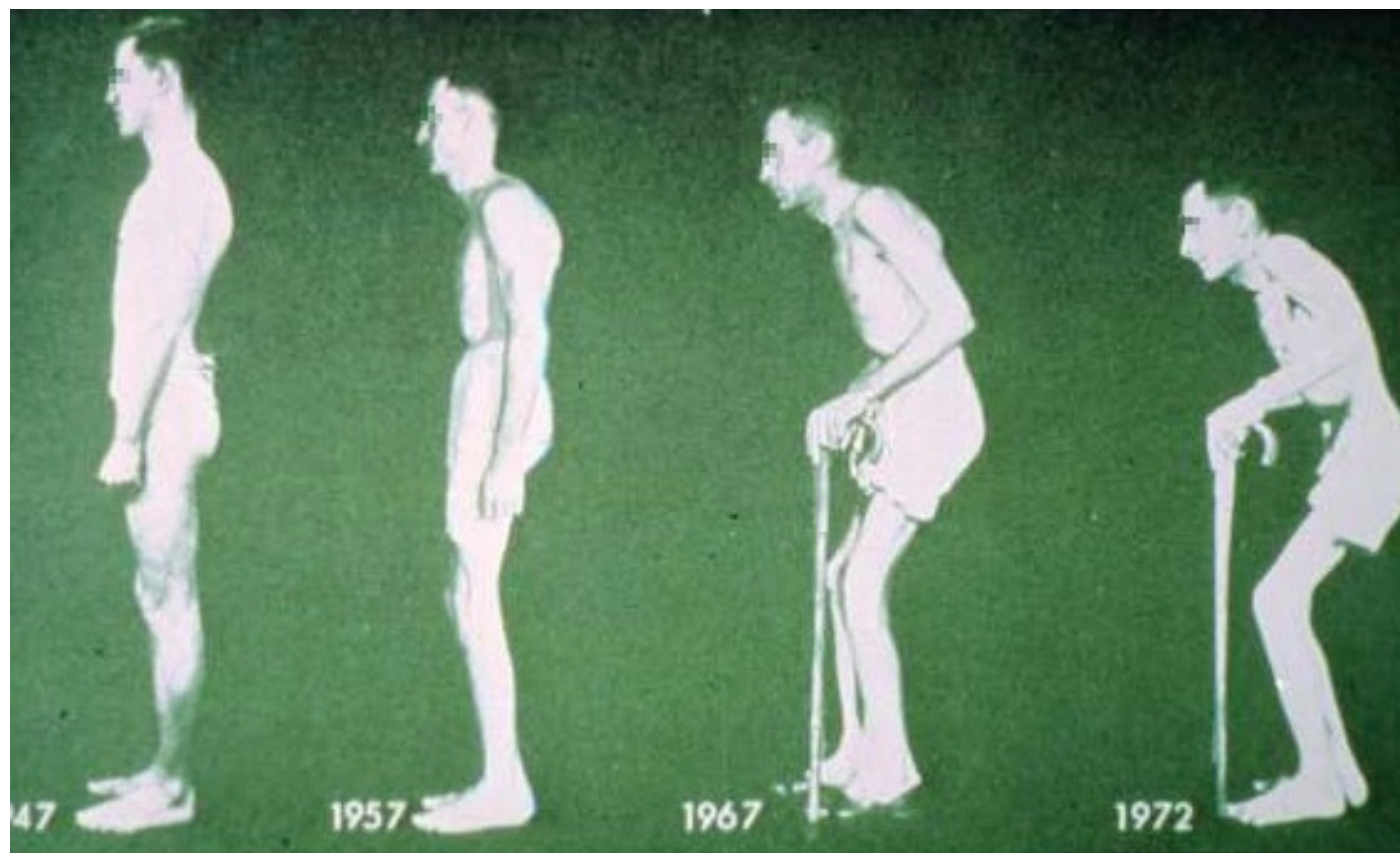
Asocjacje mogą być zwodne

- Allel 3A4 cytochromu P450 (CYP3A) i rak prostaty (bardziej zaawansowana postać w momencie diagnozy)
- CYP3A może w pływać na tempo hydroksylowania testosteronu - związek przyczynowy?
- Nie ma wpływu allelu 3A4 na kinetykę metabolizmu testosteronu
- Allel 3A4 częściej występuje u ludzi pochodzenia afrykańskiego (Afroamerykanie), niż europejskiego
- Podobne korelacje dla raka prostaty dla innych alleli częstszych w populacji afrykańskiej
- Bardziej zaawansowany rak prostaty u Afroamerykanów (przyczyny społeczne)?

Asocjacja

- Zależności funkcjonalne
 - często z allelami MHC (HLA) - związane z funkcjonowaniem układu odpornościowego
 - poszukuje się asocjacji dla bardzo intensywnie poznawanych polimorfizmów SNP w genomie człowieka
 - farmakogenetyka i farmakogenomika – polimorfizmy a działanie leków
 - CRHR1 (receptor kortykotropiny) - skuteczność terapii kortykosteroidami w leczeniu astmy
 - HLA-B27 - nadwrażliwość na skutki uboczne Abcaviru
 - HTR2A (serotonina 2A) - reakcja na środki antydepresyjne

Zesztywniające zapalenie stawów



Asocjacja - przykład

HLA-B27 i choroby autoimmunologiczne, np. zeszywniające zapalenie stawów

	Chorzy	Zdrowi
HLA-B27 +	90	1000
HLA-B27 -	10	9000

Ryzyko 8%

Ryzyko 0,11%

Ryzyko dla całej populacji ~1%

test statystyczny Fishera (Fisher exact test):

$p \approx 2 \cdot 10^{-76}$

Geny na...?

Odkryto gen odpowiedzialny za choroby psychiczne

rik, IAR

2009-11-26, ostatnia aktualizacja 2009-11-26 17:15

Naukowcy odkryli gen odpowiedzialny za skłonności do chorób psychicznych. Z badań przeprowadzonych na uniwersytecie w Edynburgu wynika, że ABCA13 jest częściowo nieaktywny u pacjentów cierpiących między innymi na schizofrenię, depresję i chorobę afektywną dwubiegunową.



Fot. Jarosław Kubalski / AG

ZOBACZ TAKŻE

- [Likwidują ośrodek. Co się stanie z chorymi?](#)
- [Z zaburzeniami psychicznymi częściej w szpitalu](#)

Gen reguluje te funkcje mózgu, które są zaburzone u osób chorych psychicznie. Naukowcy z Edynburga są przekonani, że to odkrycie jest przełomowe w rozumieniu przyczyn schorzeń psychicznych.

Może też pomóc w opracowaniu nowych leków i metod terapii. Odkrycie jest owocem pracy międzynarodowej grupy badawczej na uniwersytecie w Edynburgu. Naukowcy przebadali 2 tysiące pacjentów z zaburzeniami psychicznymi i taką samą liczbę zdrowych osób.

Alert24: Wyślij nam newsa, zdjęcia albo wideo! | SMS/MMS 605242424 | Czekają nagrody

Geny na ...?

BBC

News

Sport

Weather

Capital

TV

Radio

More...

Search BBC News

NEWS

LIVE BBC NEWS CHANNEL

News Front Page



Africa

Americas

Asia-Pacific

Europe

Middle East

South Asia

UK

England

Northern Ireland

Scotland

Wales

UK Politics

Education

Page last updated at 10:13 GMT, Thursday, 26 November 2009

E-mail this to a friend

Printable version

'Mental illness gene' discovered by Scots scientists

Scientists have discovered a gene which may help explain the causes of mental illness.

The ABCA13 gene is partially inactive in patients with severe psychological conditions such as schizophrenia, bipolar disorder and depression.

It is hoped that identifying genes which make people more likely to develop psychiatric illness may lead to new treatments being developed.



The discovery could lead to new treatments being developed

SEE ALSO

'Doctors should stop pushing drugs at depressed people'

25 Nov 09 | Health

Physical problems 'often mental'

14 Oct 09 | Health

RELATED BBC LINKS

Mental Health

Headroom

RELATED INTERNET LINKS

Edinburgh University

A w rzeczywistości...

Copyright © 2009 The American Society of Human Genetics. All rights reserved.
The American Journal of Human Genetics, 25 November 2009
doi:10.1016/j.ajhg.2009.11.003

ARTICLE

A Cytogenetic Abnormality and Rare Coding Variants Identify *ABCA13* as a Candidate Gene in Schizophrenia, Bipolar Disorder, and Depression

Abstract

Schizophrenia and bipolar disorder are leading causes of morbidity across all populations, with heritability estimates of ~80% indicating a substantial genetic component. Population genetics and genome-wide association studies suggest an overlap of genetic risk factors between these illnesses but it is unclear how this genetic component is divided between common gene polymorphisms, rare genomic copy number variants, and rare gene sequence mutations. We report evidence that the lipid transporter gene *ABCA13* is a susceptibility factor for both schizophrenia and bipolar disorder. After the initial discovery of its disruption by a chromosome abnormality in a person with schizophrenia, we resequenced *ABCA13* exons in 100 cases with schizophrenia and 100 controls. Multiple rare coding variants were identified including one nonsense and nine missense mutations and compound heterozygosity/homozygosity in six cases. Variants were genotyped in additional schizophrenia, bipolar, depression ($n > 1600$), and control ($n > 950$) cohorts and the frequency of all rare variants combined was greater than controls in schizophrenia (OR = 1.93, $p = 0.0057$) and bipolar disorder (OR = 2.71, $p = 0.00007$). The population attributable risk of these mutations was 2.2% for schizophrenia and 4.0% for bipolar disorder. In a study of 21 families of mutation carriers, we genotyped affected and unaffected relatives and found significant linkage (LOD = 4.3) of rare variants with a phenotype including schizophrenia, bipolar disorder, and major depression. These data identify a candidate gene, highlight the genetic overlap between schizophrenia, bipolar disorder, and depression, and suggest that rare coding variants may contribute significantly to risk of these disorders.

Dla porównania

	Chorzy	Zdrowi
HLA-B27 +	90	1000
HLA-B27 -	10	9000

Ryzyko 8%

Ryzyko 0,11%

test statystyczny Fishera (Fisher exact test):

$p \approx 2 \cdot 10^{-76}$

$$OR = \frac{\frac{90}{1000}}{\frac{10}{9000}} = 81$$

Ważne!!

- Asocjacja to nie jest “gen na...”!
- Czynniki ryzyka nie mają zwykle znaczenia diagnostycznego
 - Może być przydatny w diagnostyce różnicowej
- Zawsze należy analizować asocjację na tle ogólnego ryzyka w populacji, jakie są wartości bezwzględne
- Asocjacje mają zwykle niewielką przydatność dla przewidywania fenotypu pojedynczych osób - to nie jest diagnostyka

Asocjacje, testy, nadużycia

- Wiele komercyjnych testów genetycznych (DTC - *Direct To Consumer*) wykorzystuje niezrozumienie różnicy między asocjacją a mutacją sprawczą
- szkolne i wstępne kursy genetyki ograniczają się do dziedziczenia mendlowskiego

Zrobiliście listę noworocznych postanowień? Genetycy tłumaczą, dlaczego ich nie dotrzymacie

Dostępne badania:

* ciąża - pokazuje predyspozycje genetyczne w zakresie przyswajania witamin: kwasu foliowego i witaminy D. To składniki ważne dla matki i dziecka,

* otyłość - przeznaczone zarówno dla osób, które borykają się z problemem nadwagi, jak i tych, które chciałyby jej uniknąć. Sprawdza genetyczne predyspozycje do zaburzeń metabolizmu lub złych nawyków żywieniowych,

* stres - wskazane dla każdego, kto chce sprawdzić swoje predyspozycje do radzenia sobie ze stresem,

* talent - dla każdego, kto chce poznać i zrozumieć otrzymane w genach predyspozycje fizyczne, psychiczne lub intelektualne,

Choroby serca?

temu łatwo mogły określić, na co powinni uważać konkretni zawodnicy. Piłkarz Lilian Thuram zakończył karierę w 2008 roku, po tym jak badania genetyczne wykazały, że ma skłonność do chorób serca. Takie problemy miała jego matka, a brat umarł w młodym wieku na zawał serca. Stwierdził, że nie ma co ryzykować. W końcu lista piłkarzy, którzy zmarli na boisku, jest bardzo długa - mówi naukowiec. Podkreśla też, że badania genetyczne powinny być obowiązkowe. Wprowadzenie ich zaraz po urodzeniu pozwoliłoby uratować wiele osób. - Mam u siebie badania kardio, gdzie oznaczam chromosom 9, czyli rejon, który mocno wpływa na choroby serca. W Polsce każdego dnia umiera na nie 500 osób, a 100 innych ma zawał, który kończy się śmiercią. Dzięki badaniom dowiadujemy się, jak wielkie jest ryzyko wystąpienia tych chorób i jak sobie pomóc. Ludzie kupują zegarki mierzące tętno czy kroki, a nie interesują się genetyką. Przecież badania to taka wydrukowana instrukcja życia.

Choroby serca?

- Lilian Thuram choruje na rodzinną kardiomiopatię rozstrzeniową
- Forma rodzinna (genetyczna) ~50% wszystkich przypadków KR
- Obszar na chromosomie 9 - jeden z kilku powiązanych z chorobą
- Częstość KR: 1/2500 osób



Choroby serca?

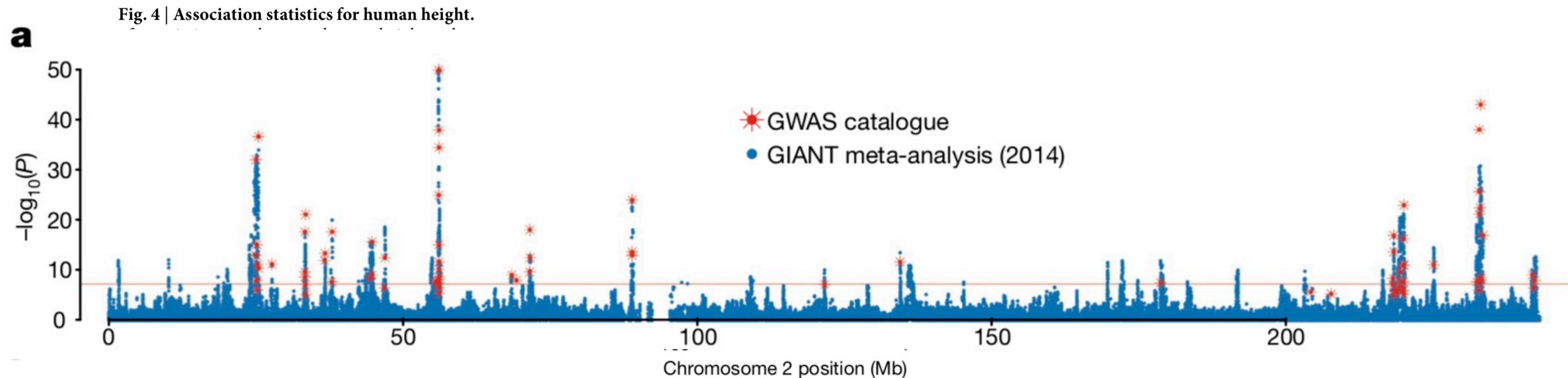
- Test chr. 9 wykrywa jeden z kilku genów stosunkowo rzadkiej genetycznej choroby serca
- Wiązanie go z zawałem (zupełnie inna choroba) i innymi częstymi chorobami serca - ????

temu łatwo mogły określić, na co powinni uważać konkretni zawodnicy. Piłkarz Lilian Thuram zakończył karierę w 2008 roku, po tym jak badania genetyczne wykazały, że ma skłonność do chorób serca. Takie problemy miała jego matka, a brat umarł w młodym wieku na zawał serca. Stwierdził, że nie ma co ryzykować. W końcu lista piłkarzy, którzy zmarli na boisku, jest bardzo długa - mówi naukowiec. Podkreśla też, że badania genetyczne powinny być obowiązkowe. Wprowadzenie ich zaraz po urodzeniu pozwoliłoby uratować wiele osób. - Mam u siebie badania kardio, gdzie oznaczam chromosom 9, czyli rejon, który mocno wpływa na choroby serca. W Polsce każdego dnia umiera na nie 500 osób, a 100 innych ma zawał, który kończy się śmiercią. Dzięki badaniom dowiadujemy się, jak wielkie jest ryzyko wystąpienia tych chorób i jak sobie pomóc. Ludzie kupują zegarki mierzące tętno czy kroki, a nie interesują się genetyką. Przecież badania to taka wydrukowana instrukcja życia.

Genetyczne czynniki ryzyka

- Badania GWAS (*Genome-Wide Association Studies*) – asocjacje na skalę genomu
- Korelacja wariantów genetycznych z ryzykiem choroby
- Znajduje się wiele korelacji, ale żadna nie może być decydującym czynnikiem
- Tajemnica brakującej odziedziczalności

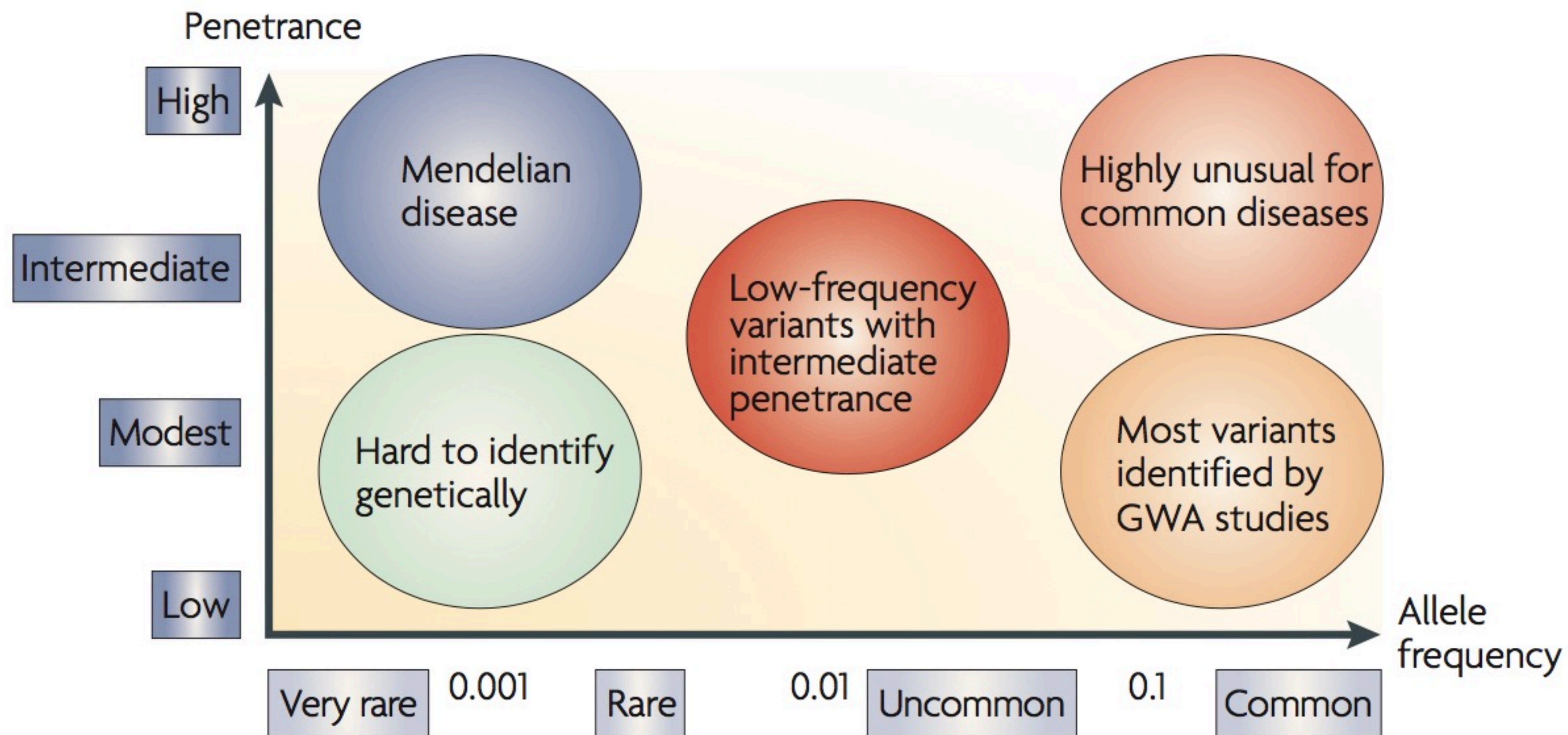
Wizualizacja GWAS - Manhattan plot



The UK Biobank resource with deep phenotyping and genomic data

Clare Bycroft^{1,13}, Colin Freeman^{1,13}, Desislava Petkova^{1,12,13}, Gavin Band¹, Lloyd T. Elliott², Kevin Sharp², Allan Motyer³, Damjan Vukcevic^{3,4}, Olivier Delaneau^{5,6,7}, Jared O'Connell⁸, Adrian Cortes^{1,9}, Samantha Welsh¹⁰, Alan Young¹¹, Mark Effingham¹⁰, Gil McVean^{1,11}, Stephen Leslie^{3,4}, Naomi Allen¹¹, Peter Donnelly^{1,2,14} & Jonathan Marchini^{1,2,14*}

Badania asocjacji



Genome-wide association studies for complex traits: consensus, uncertainty and challenges

Mark I. McCarthy^{*‡}, Gonçalo R. Abecasis[§], Lon R. Cardon^{*||}, David B. Goldstein[¶], Julian Little[#], John P. A. Ioannidis^{***†} and Joel N. Hirschhorn^{§§|||¶¶}

Nat Rev Genet, 2008 vol. 9 (5) pp. 356-369

Powtarzalność badań asocjacji

TABLE 26.2. How replicable are association studies?

Associated Gene, Phenotype	Number of Studies (Total)	Number of Studies ($p < 0.05$) Same Direction as Original Report	Number of Studies ($p < 0.05$) Opposite Direction as Original Report
<i>ABCC8</i> , type 2 diabetes ^a	9	1	1
<i>ABCC8</i>, type 2 diabetes^a	4	2	0
<i>ADD1</i>, hypertension	18	5	2
<i>APOE</i> , schizophrenia	12	0	1
<i>BLMH</i> , Alzheimer's disease	5	0	0
<i>COL1A1</i>, osteoporotic fracture	12	5	0
<i>COMT</i> , bipolar disorder	12	0	0
<i>COMT</i> , schizophrenia	9	0	1
<i>CTLA4</i>, type 1 diabetes	20	8	0
<i>DRD2</i>, schizophrenia	8	2	1
<i>DRD3</i>, schizophrenia	48	5	2
<i>GSTM1</i> , breast cancer	15	0	0
<i>GSTM1</i>, head/neck cancer	25	3	1
<i>GYS1</i> , type 2 diabetes	3	0	1
<i>HTR2A</i>, schizophrenia	28	3	1
<i>INSR</i> , type 2 diabetes	4	1	0
<i>INSR</i> , type 2 diabetes	4	0	0
<i>KCNJ11</i> , type 2 diabetes	6	0	0
<i>NTF3</i> , schizophrenia	7	0	0
<i>PON1</i>, coronary artery disease	14	5	0
<i>PPARG</i>, type 2 diabetes	14	4	1
<i>SERPINE1</i> , myocardial infarction	13	1	0
<i>SLCA1</i>, type 2 diabetes	3	2	0
<i>SLCA2</i> , type 2 diabetes	3	0	0
<i>TPH</i> , bipolar disorder	5	0	0
Total	301	47	12

Luka odziedziczalności

- “Missing heritability”
- Klasyczne badania wykazują znaczną odziedziczalność wielu cech wieloczynnikowych
- Badania asocjacyjne wykazują jedynie niewielki wzrost prawdopodobieństwa dla danego polimorfizmu (np. badania asocjacji tłumaczą tylko 5% różnic wzrostu)
- Za odziedziczalność złożonych cech wieloczynnikowych odpowiadają interakcje genetyczne wielu polimorfizmów

Zmienność liczby kopii

- Niekiedy dotyczy obszarów zawierających geny
- Ludzie o diecie wysokoskrobiowej - więcej kopii genu kodującego enzym rozkładający skrobię (amylaza, gen *AMY1*) w porównaniu z ludźmi o diecie niskoskrobiowej

Wielogenowe czynniki ryzyka

- nowa nadzieja

- GPS - *Genome-wide Polygenic Scores/PRS (Polygenic Risk Scores)*
- Łączy pojedyncze asocjacje, z których każda ma nieznaczny wpływ na fenotyp
- Analiza LDpred - uwzględnia wszystkie znane SNP, waga zależna od asocjacji

SNP	Increasing allele	Allele 1	Allele 2	Genotypic score	Correlation with trait	Weighted genotypic score
SNP 1	T	A	T	1	0.005	0.005
SNP 2	C	G	G	0	0.004	0.000
SNP 3	A	A	A	2	0.003	0.006
SNP 4	G	C	G	1	0.003	0.003
SNP 5	G	C	C	0	0.003	0.000
SNP 6	T	A	T	1	0.002	0.002
SNP 7	C	C	G	1	0.002	0.002
SNP 8	A	A	A	2	0.002	0.004
SNP 9	A	T	T	0	0.001	0.000
SNP 10	C	C	G	1	0.001	0.001
Polygenic score				9		0.023

The new genetics of intelligence

Robert Plomin¹ and Sophie von Stumm²

NATURE REVIEWS | GENETICS

VOLUME 19 | MARCH 2018 | 159

Wielogenowe czynniki ryzyka

- Współczynnik ryzyka wczesnej choroby wieńcowej na podstawie analizy 182 znanych zasocjowanych polimorfizmów
- Jednoznaczna predykcja tylko dla skrajnych wartości
- To nadal nie jest diagnostyka, ale może się przydać w identyfikowaniu grup największego ryzyka

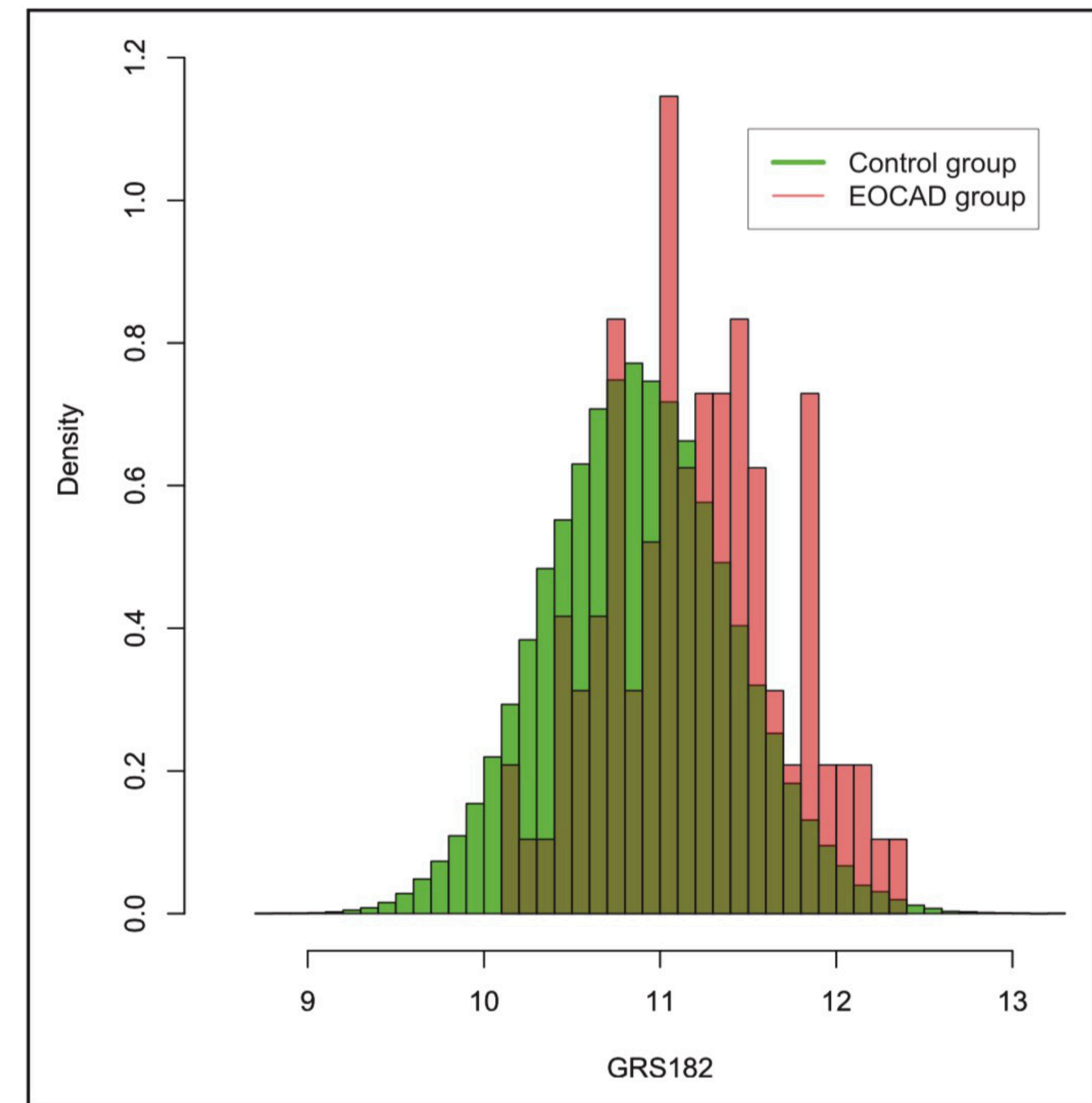


Figure 1. Distribution of genetic risk score based on the presence of 182 independent variants associated with coronary artery disease (GRS182) in the UK Biobank cohort according to early-onset coronary artery disease (EOCAD) status.

Choroba wieńcowa

- CAD (*Coronary Arterial Disease*)
- ~60 000 chorych, ~ 120 000 kontroli (UK Biobank)
- 6 630 150 SNP

Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations

Amit V. Khera^{1,2,3,4,5}, Mark Chaffin^{4,5}, Krishna G. Aragam^{1,2,3,4}, Mary E. Haas⁴, Carolina Roselli⁴, Seung Hoan Choi⁴, Pradeep Natarajan^{2,3,4}, Eric S. Lander⁴, Steven A. Lubitz^{2,3,4}, Patrick T. Ellinor^{2,3,4} and Sekar Kathiresan^{1,2,3,4*}

NATURE GENETICS | VOL 50 | SEPTEMBER 2018 | 1219-1224

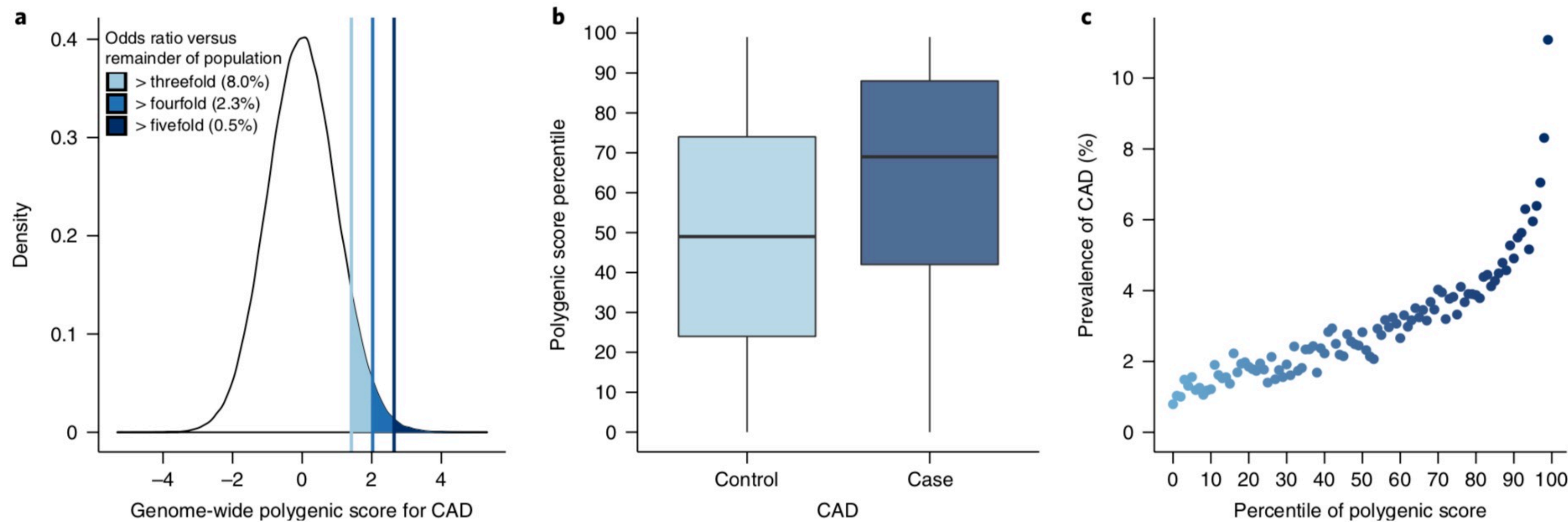


Fig. 2 | Risk for CAD according to GPS. **a**, Distribution of GPS_{CAD} in the UK Biobank testing dataset ($n = 288,978$). The x axis represents GPS_{CAD} , with values scaled to a mean of 0 and a standard deviation of 1 to facilitate interpretation. Shading reflects the proportion of the population with three-, four-, and

Stratyfikacja ryzyka

- Bardzo znaczny wzrost ryzyka u stosunkowo nielicznej grupy o wysokim GPS/PRS

High GPS definition	Reference group	Odds ratio	95% CI
CAD			
Top 20% of distribution	Remaining 80%	2.55	2.43-2.67
Top 10% of distribution	Remaining 90%	2.89	2.74-3.05
Top 5% of distribution	Remaining 95%	3.34	3.12-3.58
Top 1% of distribution	Remaining 99%	4.83	4.25-5.46
Top 0.5% of distribution	Remaining 99.5%	5.17	4.34-6.12

Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations

Amit V. Khera^{1,2,3,4,5}, Mark Chaffin^{4,5}, Krishna G. Aragam^{1,2,3,4}, Mary E. Haas⁴, Carolina Roselli⁴, Seung Hoan Choi⁴, Pradeep Natarajan^{2,3,4}, Eric S. Lander⁴, Steven A. Lubitz^{2,3,4}, Patrick T. Ellinor^{2,3,4} and Sekar Kathiresan^{1,2,3,4*}

NATURE GENETICS | VOL 50 | SEPTEMBER 2018 | 1219-1224

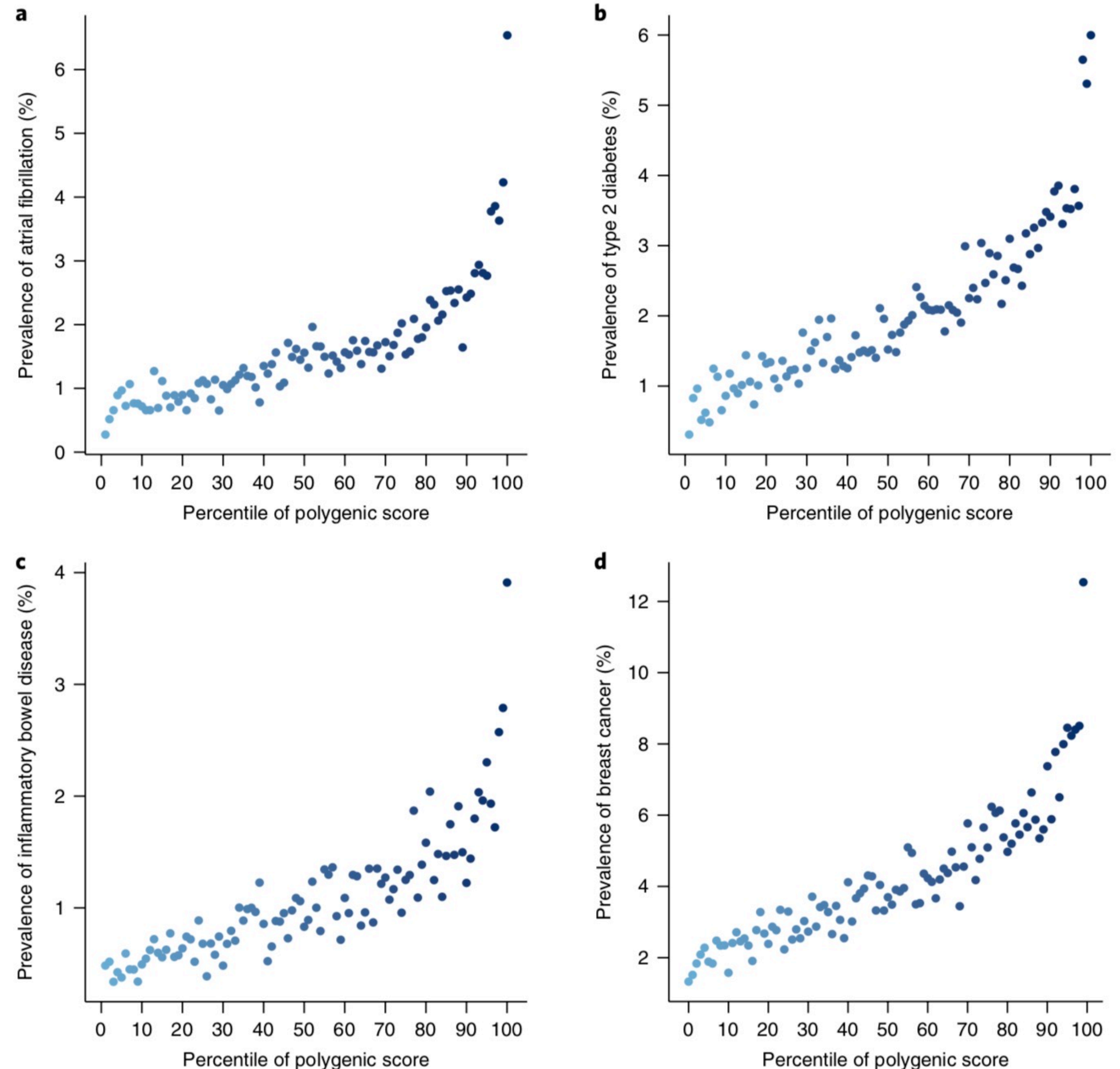
Inne choroby

- Dla 4 innych badanych chorób podobny przebieg krzywej
- (a) migotanie przedsionków, (b) cukrzyca typu 2, (c) nieswoiste zapalenie jelit, (d) rak piersi
- Dla diagnostyki i profilaktyki istotne osoby z górnych percentyli rozkładu GPS/PRS

Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations

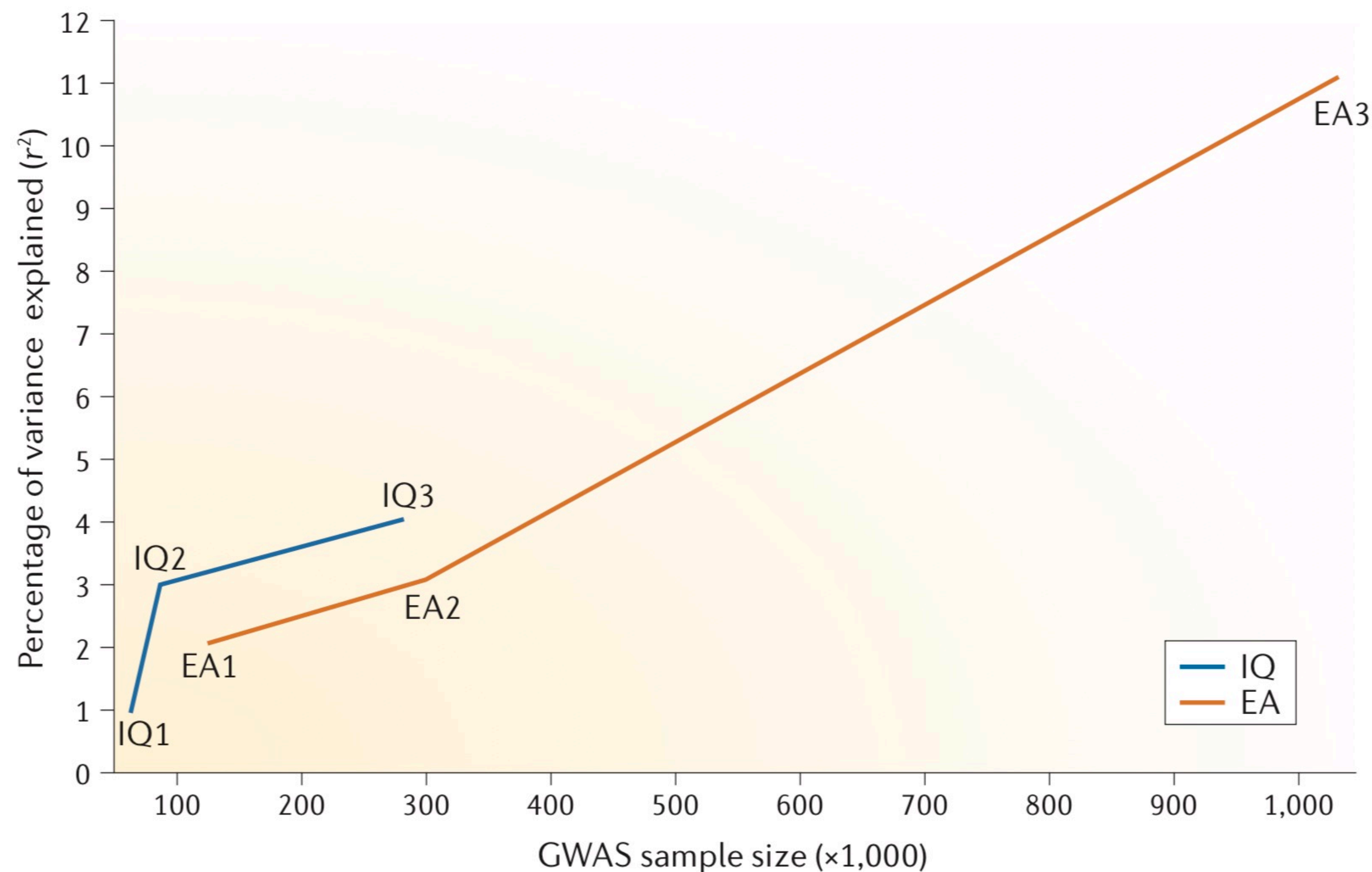
Amit V. Khera^{1,2,3,4,5}, Mark Chaffin^{4,5}, Krishna G. Aragam^{1,2,3,4}, Mary E. Haas⁴, Carolina Roselli⁴, Seung Hoan Choi⁴, Pradeep Natarajan^{2,3,4}, Eric S. Lander⁴, Steven A. Lubitz^{2,3,4}, Patrick T. Ellinor^{2,3,4} and Sekar Kathiresan^{1,2,3,4*}

NATURE GENETICS | VOL 50 | SEPTEMBER 2018 | 1219–1224



Wielkość próby

- Analizy typu GWAS/GPS wyjaśniają zmienność dopiero przy zebraniu danych z bardzo licznych grup
- Analiza dziedziczenia wieloczynnikowego wymaga dużych prób!
- W odróżnieniu od jednogenowych mutacji sprawczych (gdzie czasami wystarcza jedna rodzina)



The new genetics of intelligence

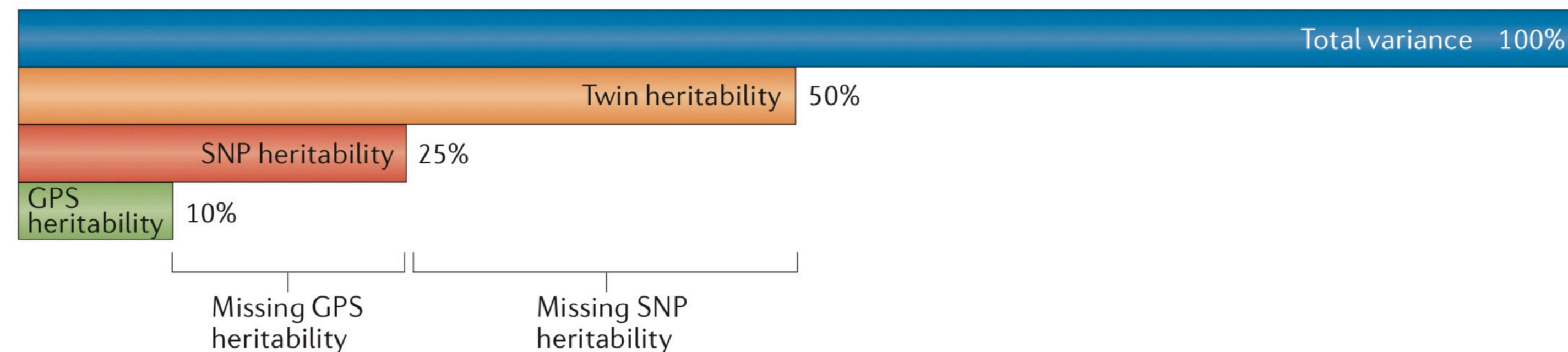
Robert Plomin¹ and Sophie von Stumm²

NATURE REVIEWS | GENETICS

VOLUME 19 | MARCH 2018 | 159

Nadal brakująca odziedziczalność

- Dla inteligencji odziedziczalność szacowana na ~50%
- Odziedziczalność zależna od SNP ~25%
- maksimum, które można wyjaśnić GPS
- Obecnie ~10%



The new genetics of intelligence

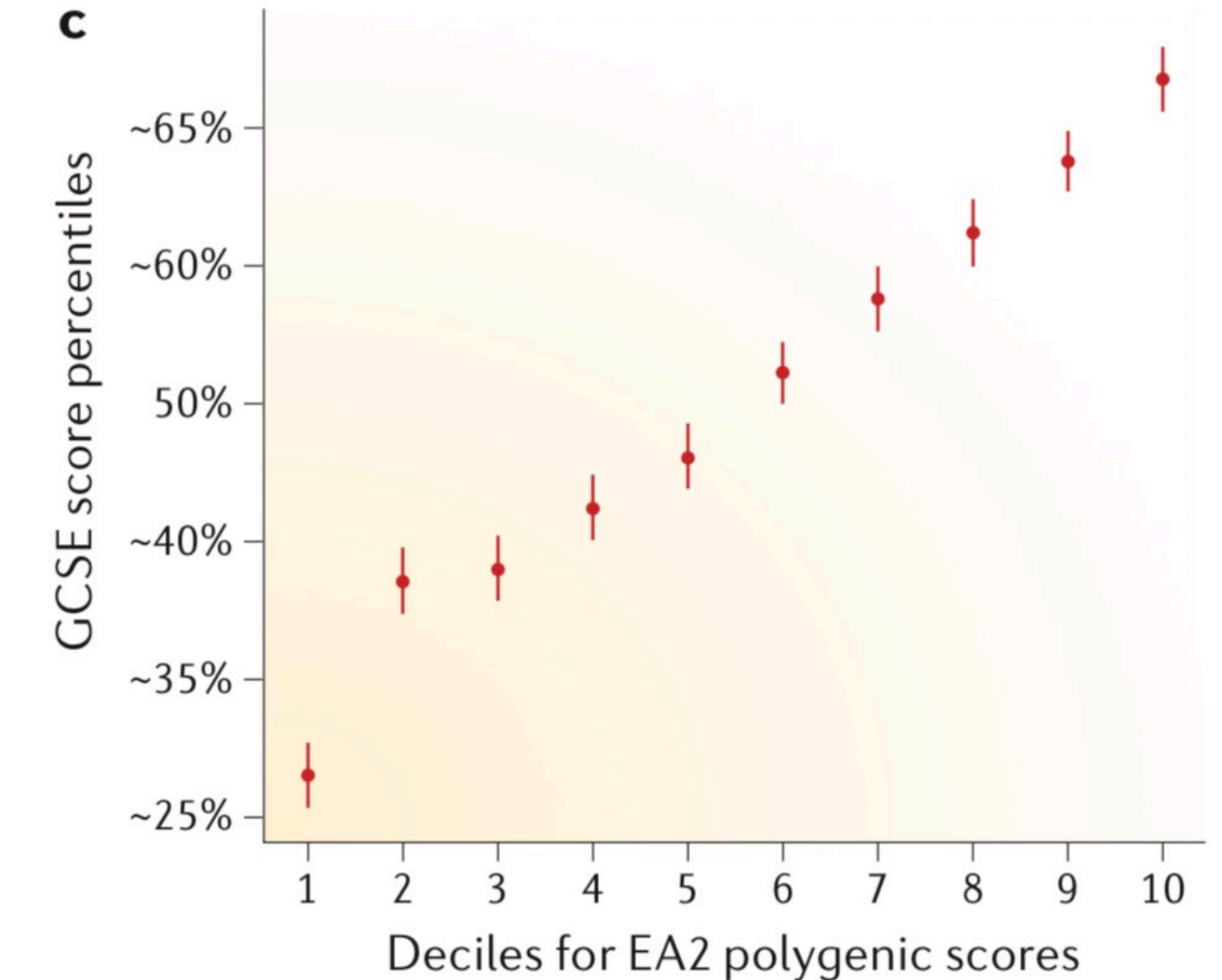
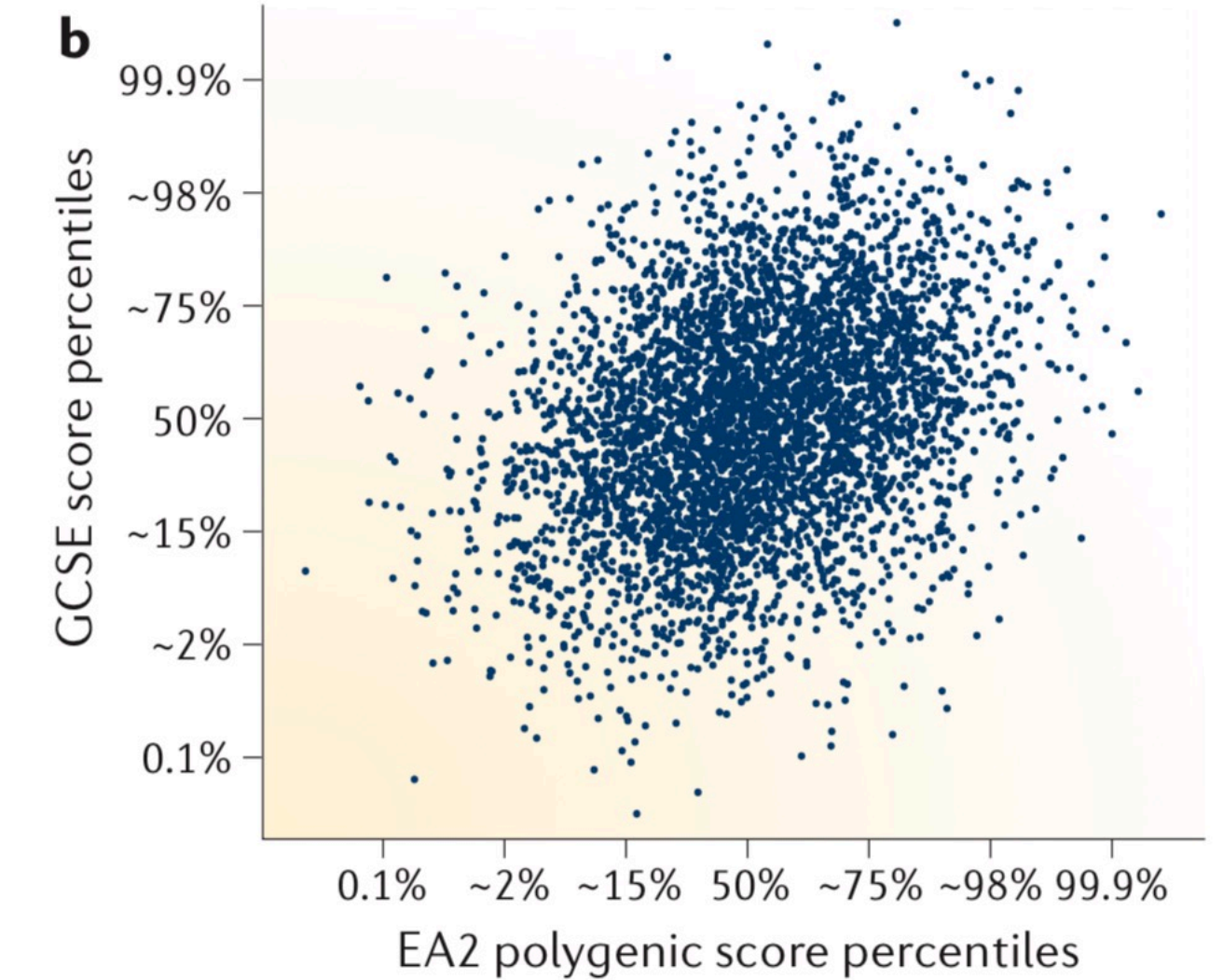
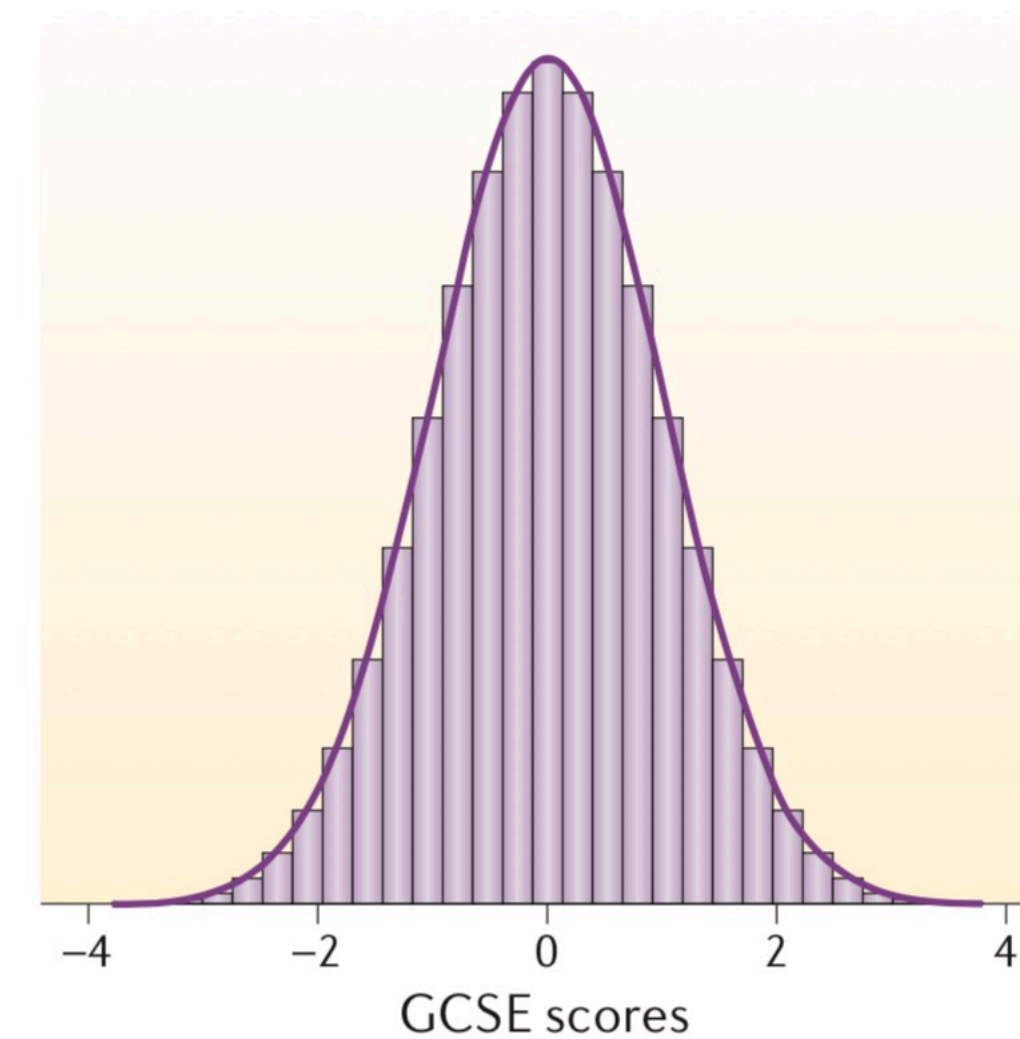
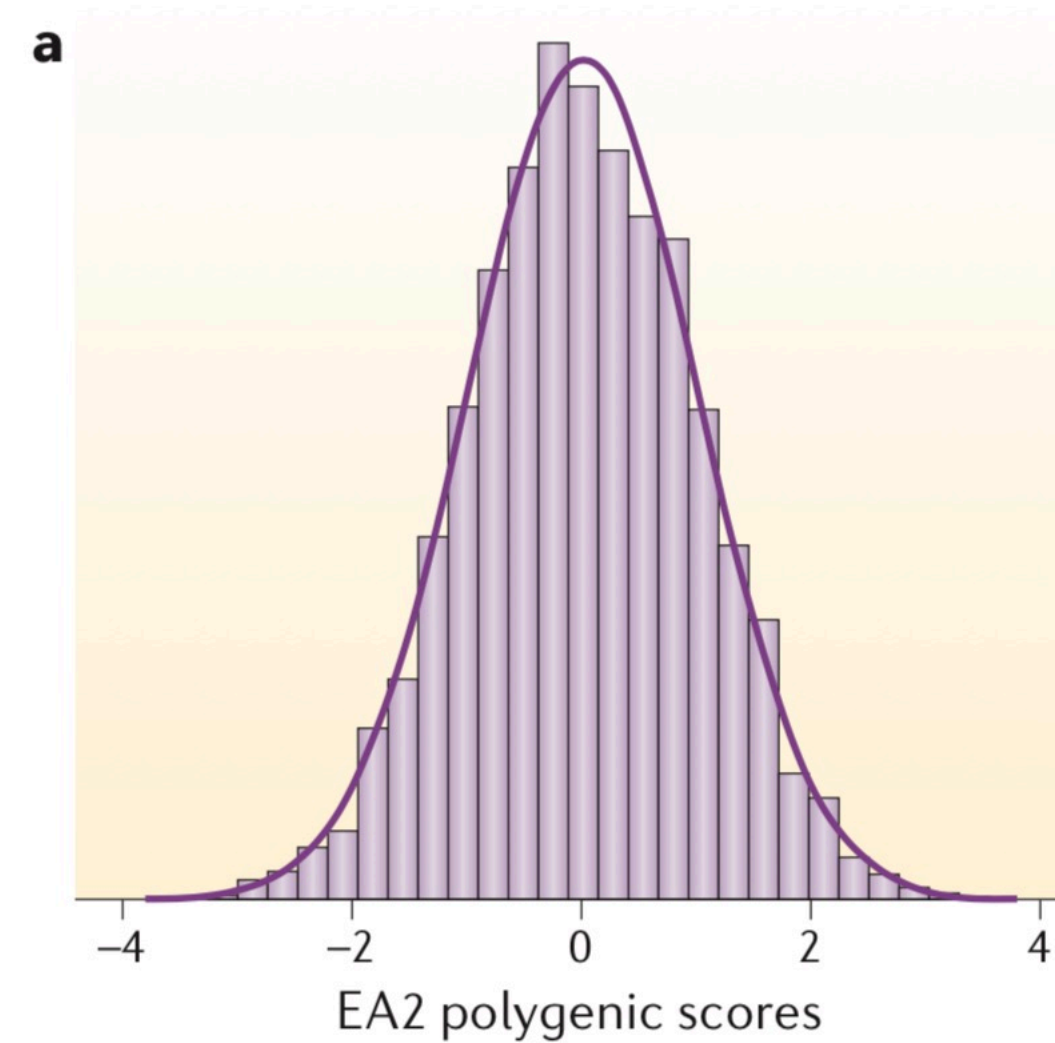
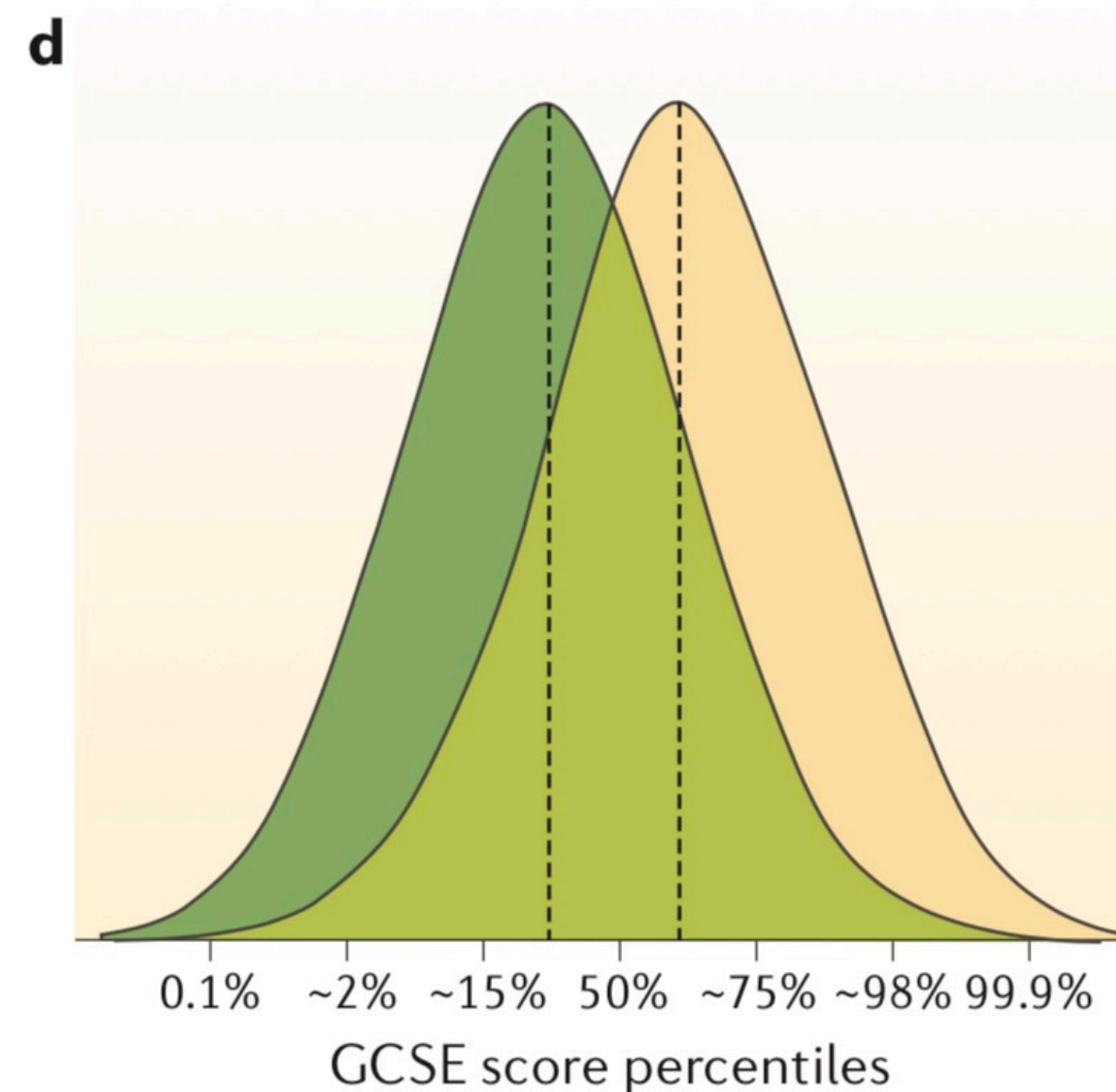
Robert Plomin¹ and Sophie von Stumm²

NATURE REVIEWS | GENETICS

VOLUME 19 | MARCH 2018 | 159

To nadal nie jest diagnostyka

- Zależność wyników w nauce (test GCSE) od GPS dla inteligencji
- Nakładanie się rozkładów GCSE dla najwyższego i najniższego decyla GPS:



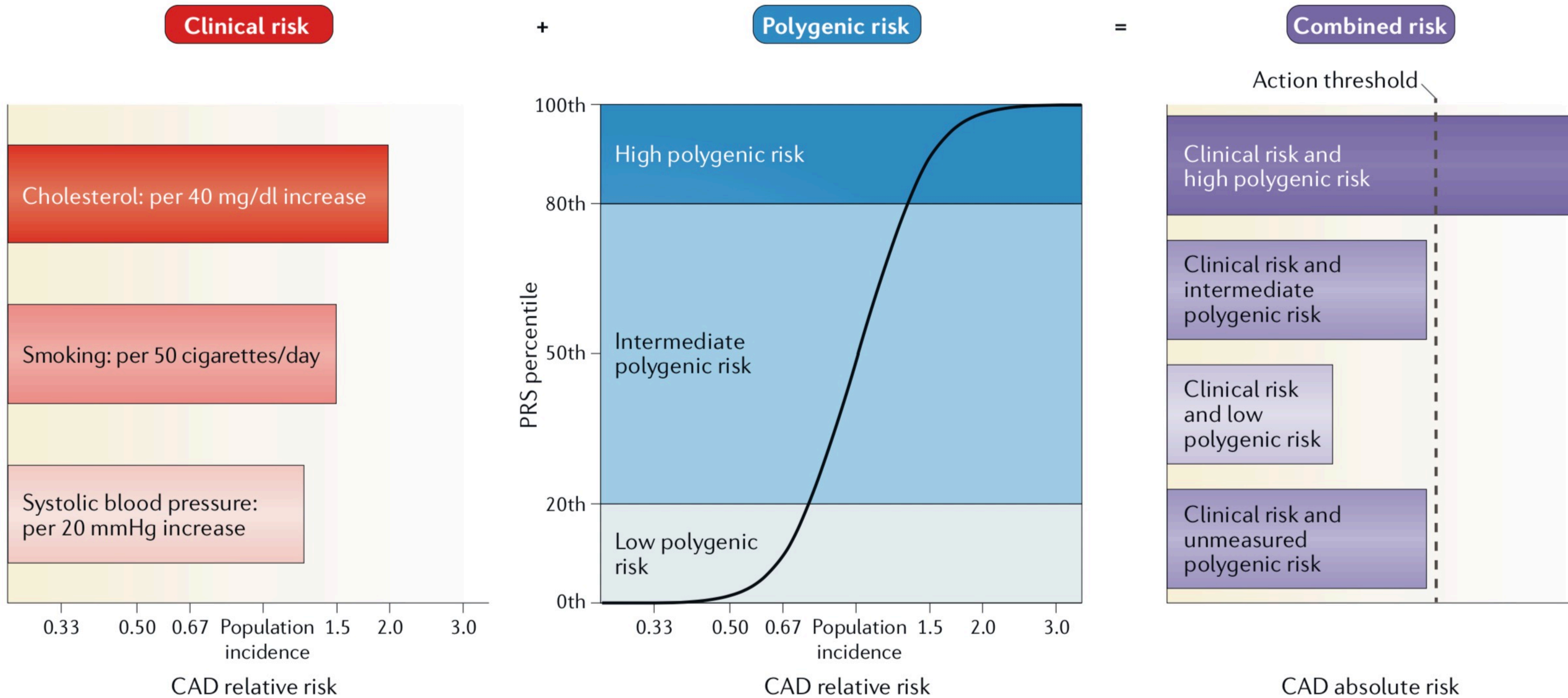
The new genetics of intelligence

Robert Plomin¹ and Sophie von Stumm²

NATURE REVIEWS | GENETICS

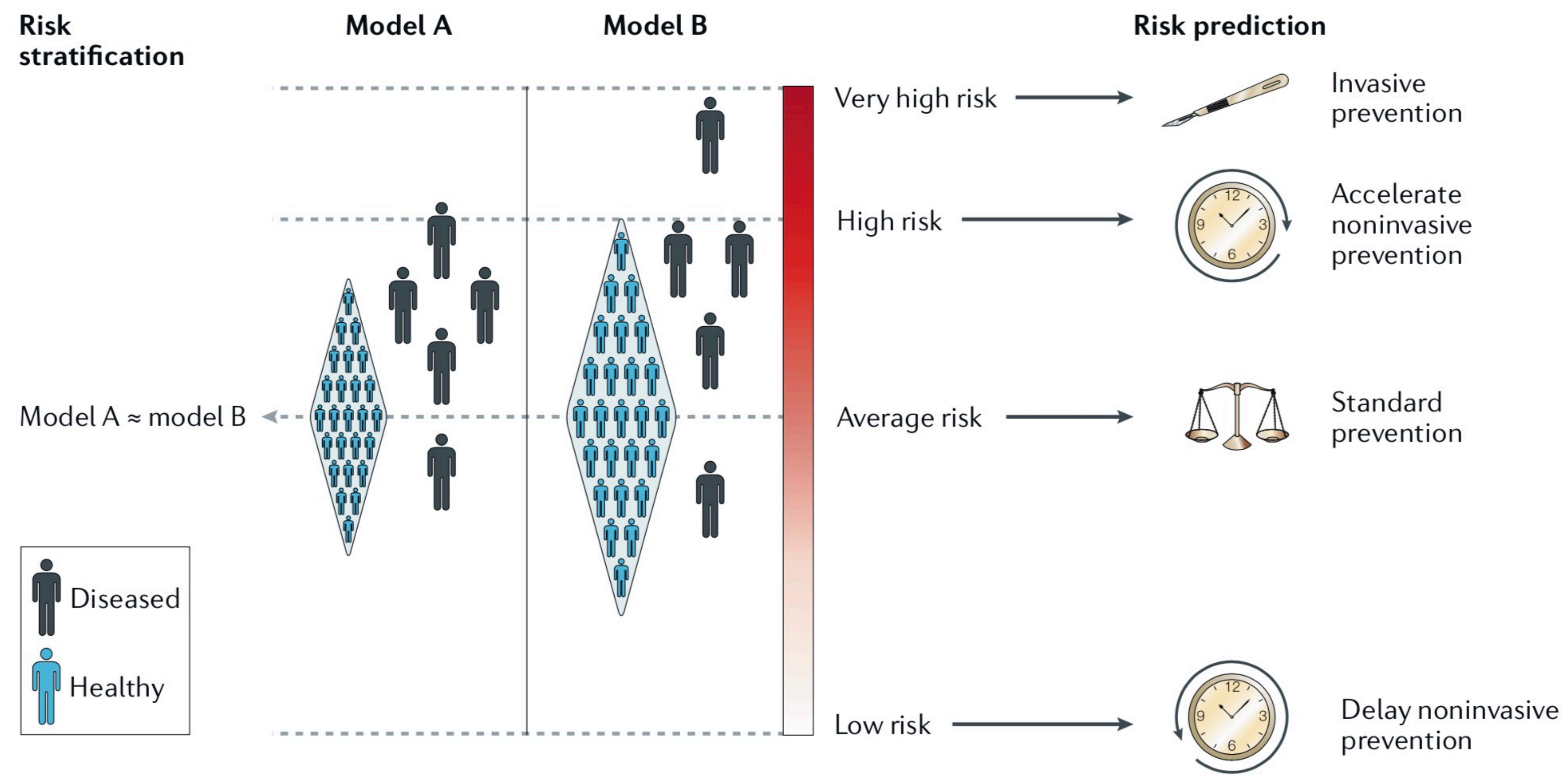
VOLUME 19 | MARCH 2018 | 159

Ryzyko środowiskowe i genetyczne



The personal and clinical utility of polygenic risk scores

Stratyfikacja ryzyka



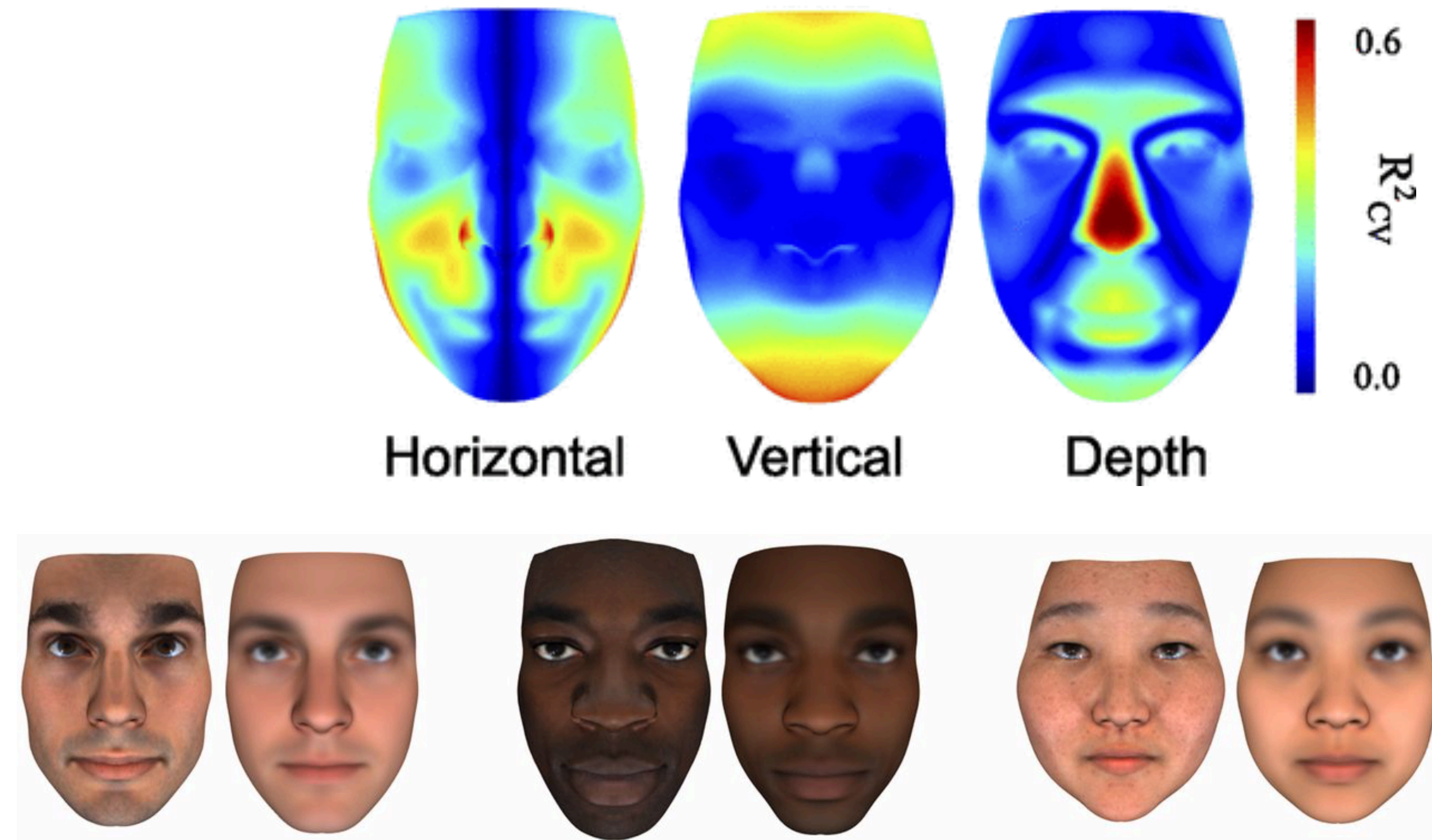
Względne ryzyko podobne w modelu A i B, ale model B lepiej stratyfikuje

PRS w badaniach przesiewowych

- Zastosowanie PRS obok standardowych metod może poprawić dodatnią wartość predykcyjną badań przesiewowych
- Np. dodatnia wartość predykcyjna testu PSA dla raka prostaty w populacji tp ~12,5% (duża częstość wyników fałszywie pozytywnych)
- Test PSA na grupie górnych 5% PRS ma dodatnią wartość predykcyjną ~25%

Sztuczna inteligencja

- Podejście wykorzystujące uczenie maszynowe (*machine learning*) na dużych zbiorach danych
- Przykład: predykcja cech wyglądu twarzy na podstawie sekwencji całogenomowych
- Czy to koniec nauki?
 - predykcja działa, ale nie rozumiemy więcej, niż dotychczas



Identification of individuals by trait prediction using whole-genome sequencing data

Christoph Lippert^{a,1}, Riccardo Sabatini^a, M. Cyrus Maher^a, Eun Yong Kang^a, Seunghak Lee^a, Okan Arikan^a, Alena Harley^a, Axel Bernal^a, Peter Garst^a, Victor Lavrenko^a, Ken Yocum^a, Theodore Wong^a, Mingfu Zhu^a, Wen-Yun Yang^a, Chris Chang^a, Tim Lu^b, Charlie W. H. Lee^b, Barry Hicks^a, Smriti Ramakrishnan^a, Haibao Tang^a, Chao Xie^c, Jason Piper^c, Suzanne Brewerton^c, Yaron Turpaz^{b,c}, Amalio Telenti^b, Rhonda K. Roby^{b,d,2}, Franz J. Och^a, and J. Craig Venter^{b,d,1}