BIOTECH 201

DECIPHERING OUR GENOMIC MESSAGES

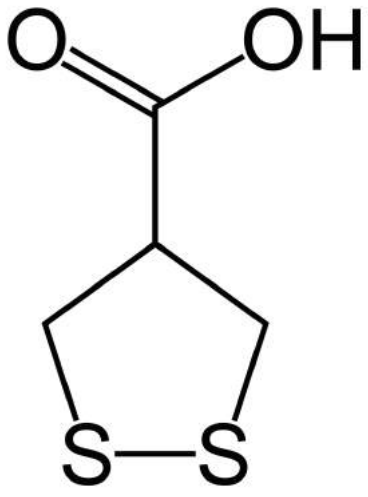
Week 3 - February 21, 2017

HUDSONALPHA
INSTITUTE FOR BIOTECHNOLOGY

Asparagus

Asparagus officinalis

asparagusic acid - $C_4H_6O_2S_2$



Asparagus anosmia

Asparagus officinalis

Dante Alighieri's portrait by Sandro Botticell

RESEARCH

OPEN ACCESS

Sniffing out significant "Pee values": genome wide association study of asparagus anosmia

Sarah C Markt,[1] Elizabeth Nuttall,[1] Constance Turman,[4] Jennifer Sinnott,[1,5] Eric B Rimm,[1,2,6] Ethan Ecsedy,[7] Robert H Unger,[1] Katja Fall,[1,8,9] Stephen Finn,[10] Majken K Jensen,[2,6] Jennifer R Rider,[1,11] Peter Kraft,[1,3,4] Lorelei A Mucci[1,6,9]

[1]Department of Epidemiology, Harvard TH Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA
[2]Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA, USA
[3]Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA, USA

ABSTRACT

OBJECTIVE
To determine the inherited factors associated with the ability to smell asparagus metabolites in urine.

DESIGN
Genome wide association study.

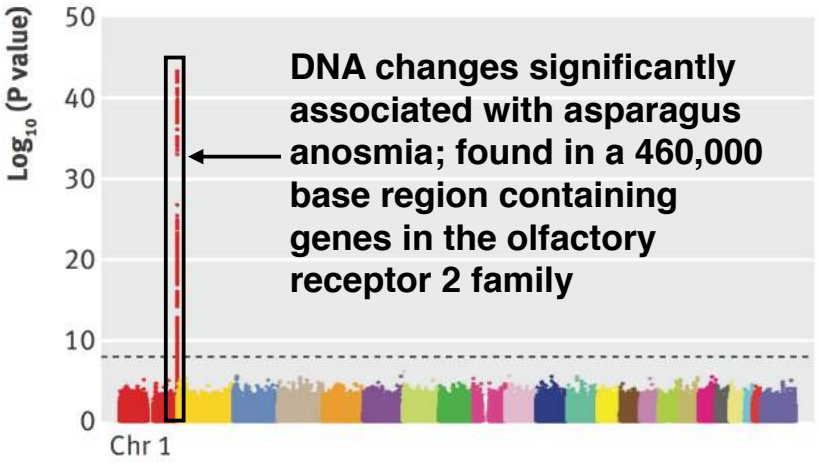
SETTING
Nurses' Health Study and Health Professionals

Introduction
In 1781 Benjamin Franklin remarked, "a few stems of asparagus eaten, shall give our urine a disagreeable odour."[1-3] The consequence of asparagus consumption has been a topic of both public and private discussion, with Proust's observation of asparagus spears perhaps the most poetic, "they played . . . at transforming my

Among 6909 participants, 39.8% (n=2748) strongly agreed that they could perceive a distinct odor in their urine after eating asparagus and 60.3% (n=4161) said they could not

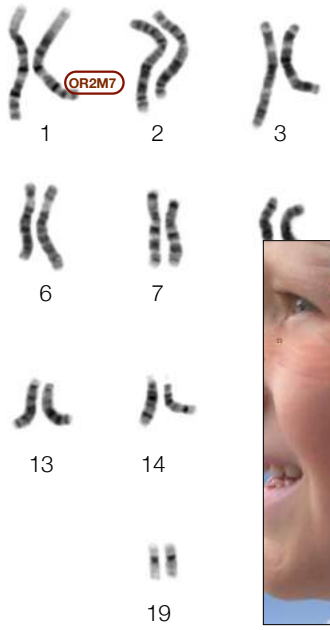
Asparagus anosmia

Asparagus officinalis



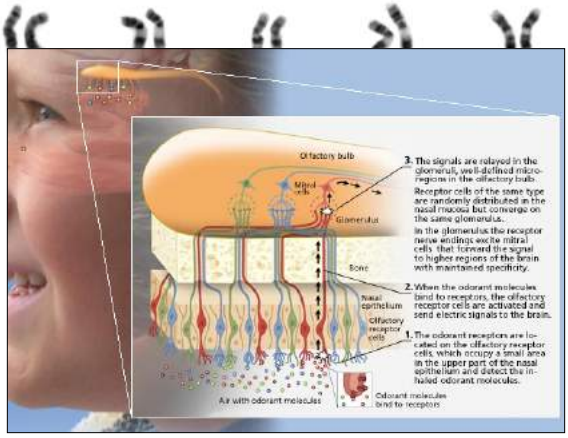
"Future replication studies are necessary before considering targeted therapies to help anosmic people discover what they are missing."

how about in humans?



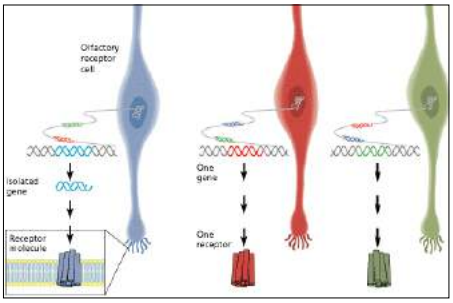
http://www.genome.gov/glossary/resources/karyotype.pdf

presumably, the DNA change alters the function of the olfactory receptor, rendering the aroma undistinguishable



OR2M7 - top candidate gene in this region

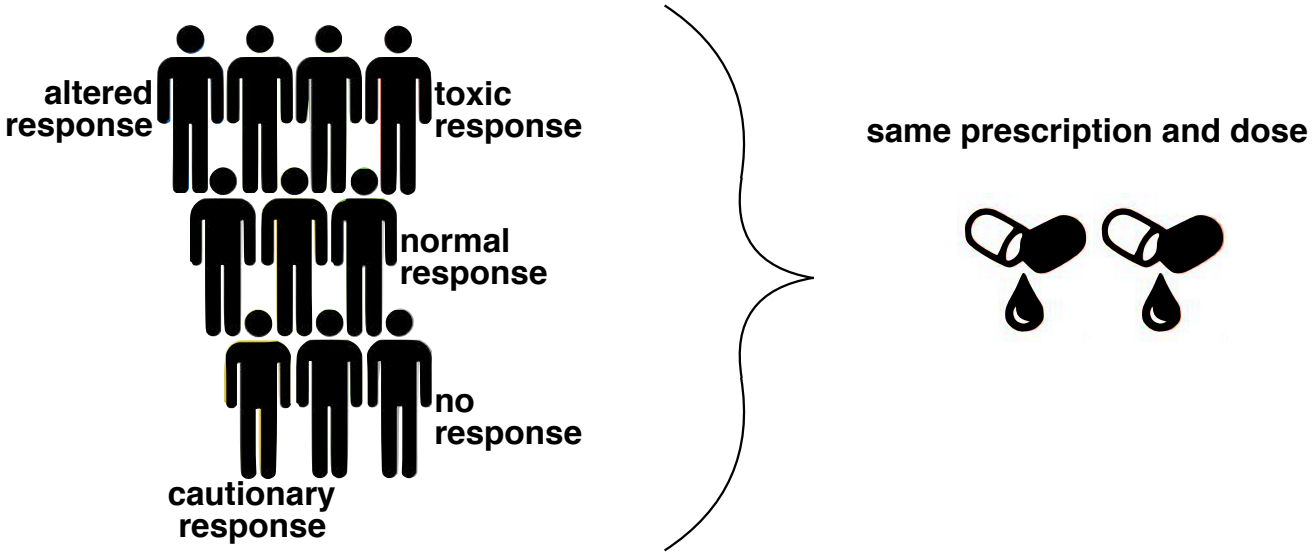
- **gene is 938 nucleotides long**
- **“olfactory receptor” gene**
- **encodes a 312 amino acid protein**



http://olfactory-system.blogspot.com/2007/02/odorant-receptors-organization-of.html

Pharmacogenomic Testing

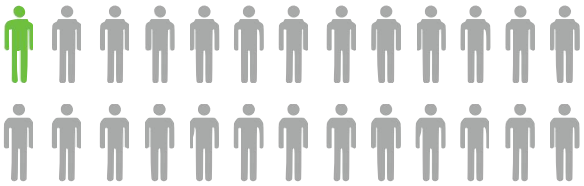
Traditional drug treatment approach



Drug Effectiveness

1. ABILIFY (aripiprazole)
Schizophrenia

2. NEXIUM (esomeprazole)
Heartburn



3. HUMIRA (adalimumab)
Arthritis

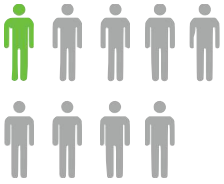
4. CRESTOR (rosuvastatin)
High cholesterol

For every person they do help (green), the ten highest grossing drugs in the U.S. fail to improve the conditions of even more people (grey)

go.nature.com/4dr78f

Drug Effectiveness

5. CYMBALTA (duloxetine)
Depression



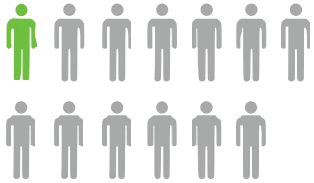
6. ADVAIR DISKUS (fluticasone propionate)
Asthma

7. ENBREL (entanercept)
Proriasis

8. REMICADE (infliximab)
Crohn's disease

9. COPAXONE (glatiramer acetate)
Multiple sclerosis

10. NEULASTA (pegfilgrastim)
Neutropenia

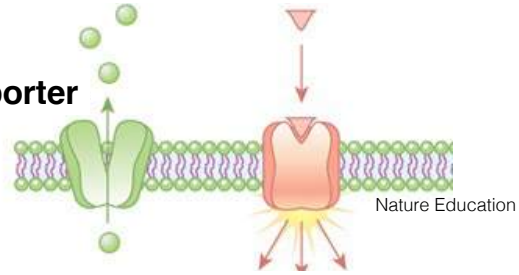


For every person they do help (green), the ten highest grossing drugs in the U.S. fail to improve the conditions of even more people (grey)

go.nature.com/4dr78f

Genetic Influences

drug transporter

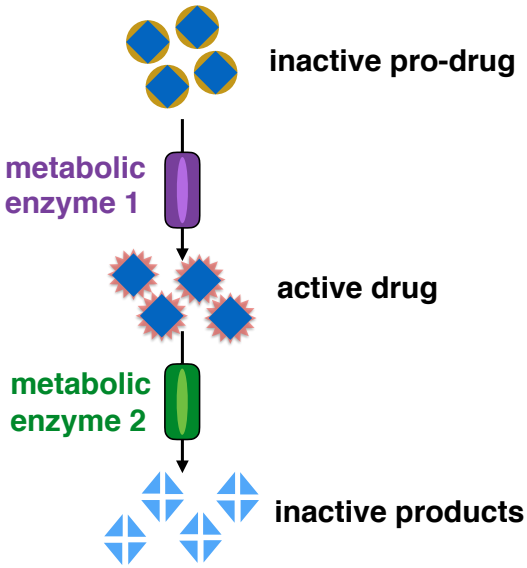


drug receptor

variation in the genes that encode transporters, receptors and metabolic enzymes impacts the concentration of active drugs in the body

the duration and intensity of a drug depends on the rate of metabolic conversion

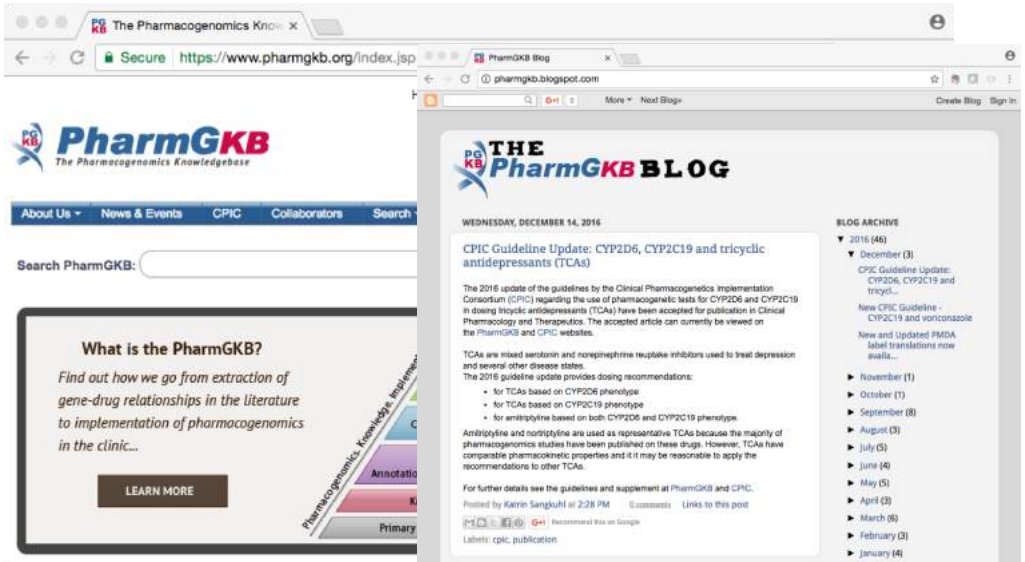
drug metabolism

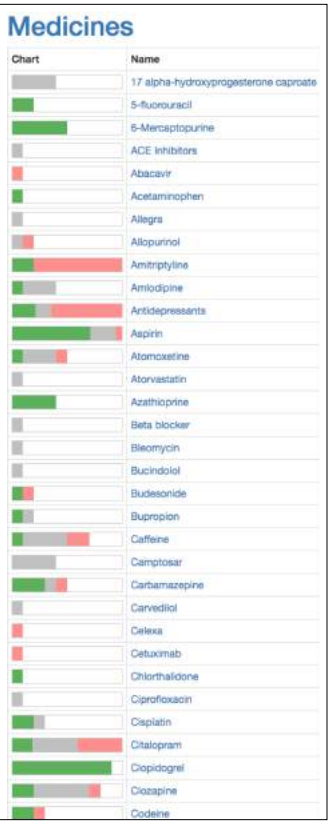


Pharmacogenomic Testing

176 FDA-approved drugs with pharmacogenomic information in the labeling

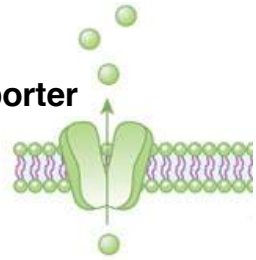
as of 11/07/16





Genetic Influences

drug transporter



The ABCB1 gene (on chromosome 7) encodes a transporter protein that moves certain molecules across the blood-brain barrier

there are multiple DNA variants in this gene - several are associated with a better response to certain antidepressant medications

RS4148739 - DNA variant A or G

A = less effective G = greater transport

frequency across world populations

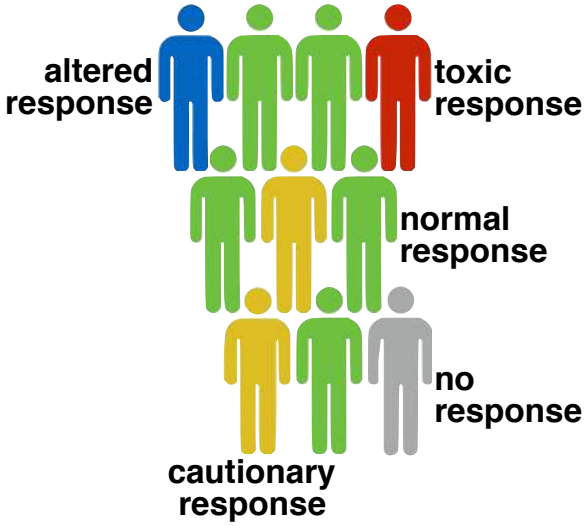
A,A genotype

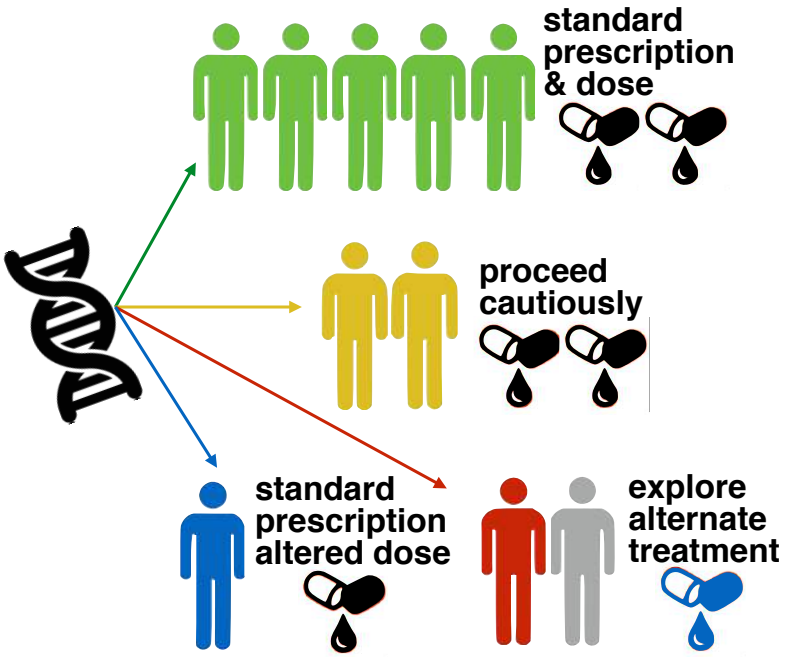
A,G genotype

G,G genotype

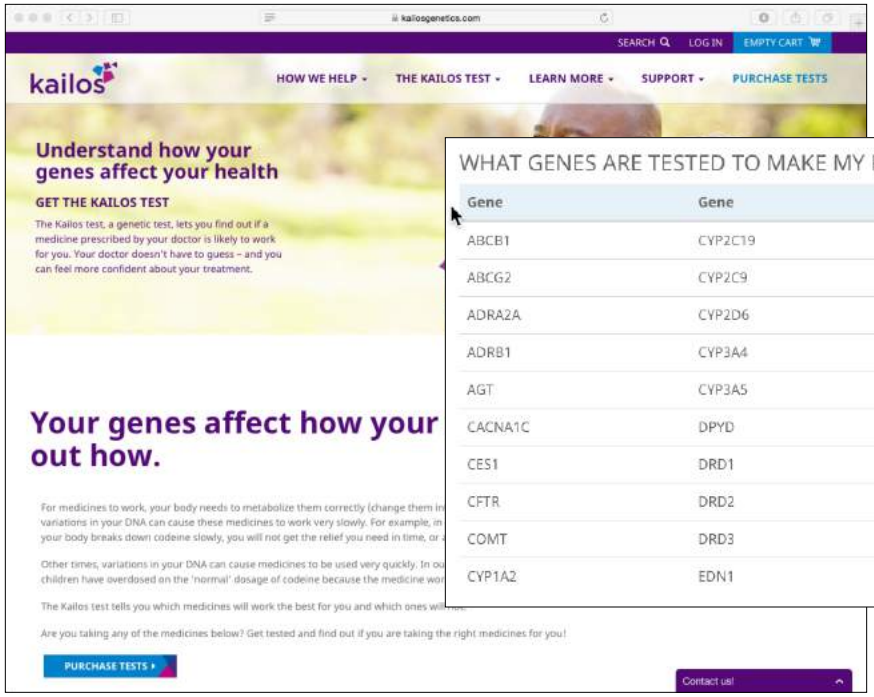
A,G and G,G individuals are 7X more likely to respond positively to certain antidepressants

Pharmacogenomic testing





Pharmacogenomic Testing



WHAT GENES ARE TESTED TO MAKE MY PGXCOMPLETE REPORT?

Gene	Gene	Gene	Gene
ABCB1	CYP2C19	F2	MTHFR
ABCG2	CYP2C9	F5	NR1H3
ADRA2A	CYP2D6	GNB3	OPRM1
ADRB1	CYP3A4	GRIK4	RYR1
AGT	CYP3A5	HTR1A	SLC6A2
CACNA1C	DPYD	HTR2A	SLCO1B1
CES1	DRD1	HTR2C	TPMT
CFTR	DRD2	IFNL3	VKORC1
COMT	DRD3	KCNIP1	
CYP1A2	EDN1	LDLR	

Cancer Testing

As a reminder

- **Cancer is a genetic disease - caused by mutations in the genes that control cell growth and repair pathways.**
- **However, these mutations are primarily acquired, rather than inherited.**
- **5-10% of breast, ovarian, colon and a few other cancers may be heavily influenced by inherited mutations, passed from generation to generation (family cancer syndromes)**

Want more details? Check out the 2016 Biotech201 series on cancer at http://hudsonalpha.org/biotech-201

Two types of “cancer” genetic tests

1. predisposition testing

- **Did you inherit a mutation that gives you a higher lifetime risk for cancer?**

2. tumor genetic analysis

- **What mutations are driving the development and progression of the existing cancer**

Two types of “cancer” genetic tests

1. predisposition testing

- **Did you inherit a mutation that gives you a higher lifetime risk for cancer?**

the impact on cancer risk varies greatly from gene to gene - some raise lifetime risk to 70-80%, others only to approximately 15%

Genes associated with inherited breast/ovarian cancer

ATM	NBN
BARD1	PALB2
BRCA1	PMS2
BRCA2	PTEN
BRIP1	RAD50
CDH1	RAD51C
CHEK2	RAD51D
EPCAM	RINT1
MLH1	STK11
MRE11A	TP53
MSH2	XRCC2
MSH6	

Two types of "cancer" genetic tests

1. predisposition testing

- **Did you inherit a mutation that gives you a higher lifetime risk for cancer?**

Two types of "cancer" genetic tests

Information is **POWER**

free genetic testing for cancer risk

- community initiative - tests for predisposition to breast/ovarian cancer
- 23 known predisposition genes
- 5 county N. AL region
- free for 30 year olds; $129 for other adults
- over 1,500 individuals tested to date

REDSTONE FEDERAL CREDIT UNION

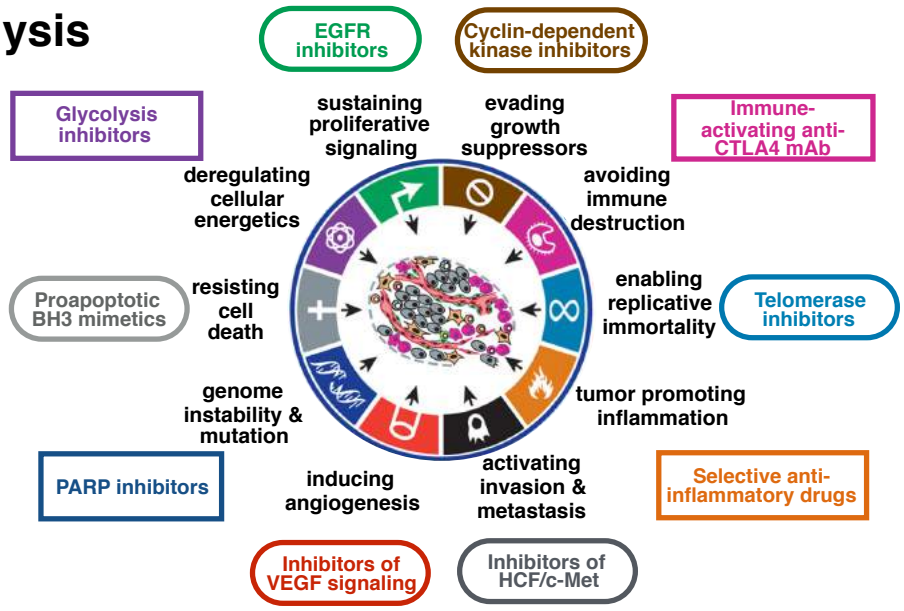
HUDSONALPHA INSTITUTE FOR BIOTECHNOLOGY

KAILOS GENETICS

Two types of "cancer" genetic tests

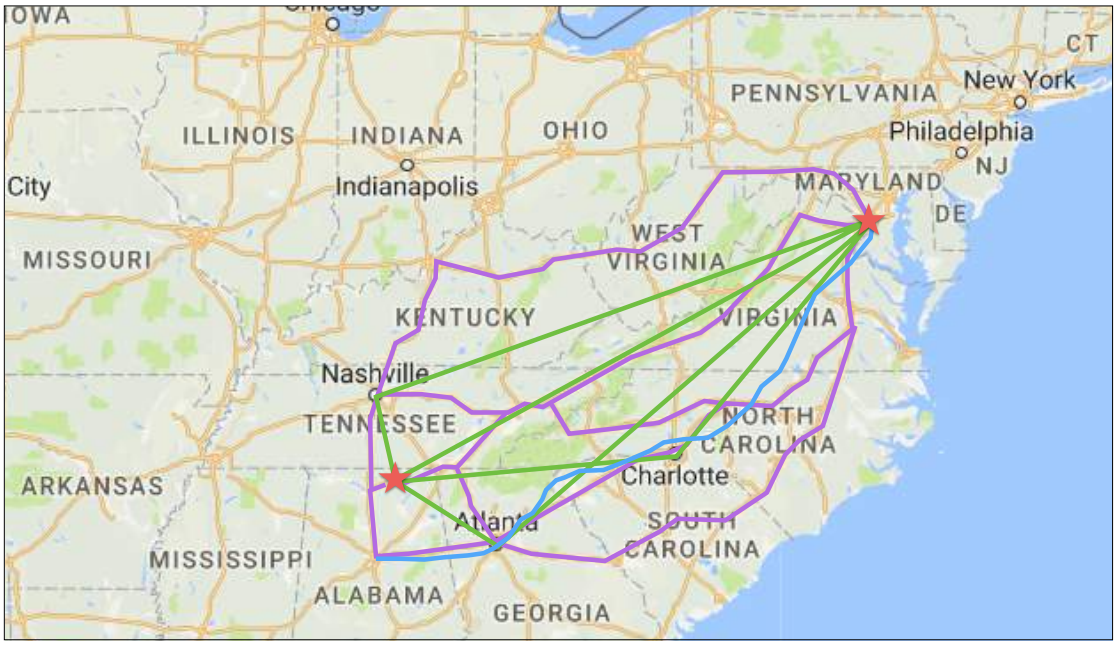
2. tumor genetic analysis

- **What mutations are driving the development and progression of the existing cancer**
- **can shape diagnosis, prognosis and therapeutic decision-making**

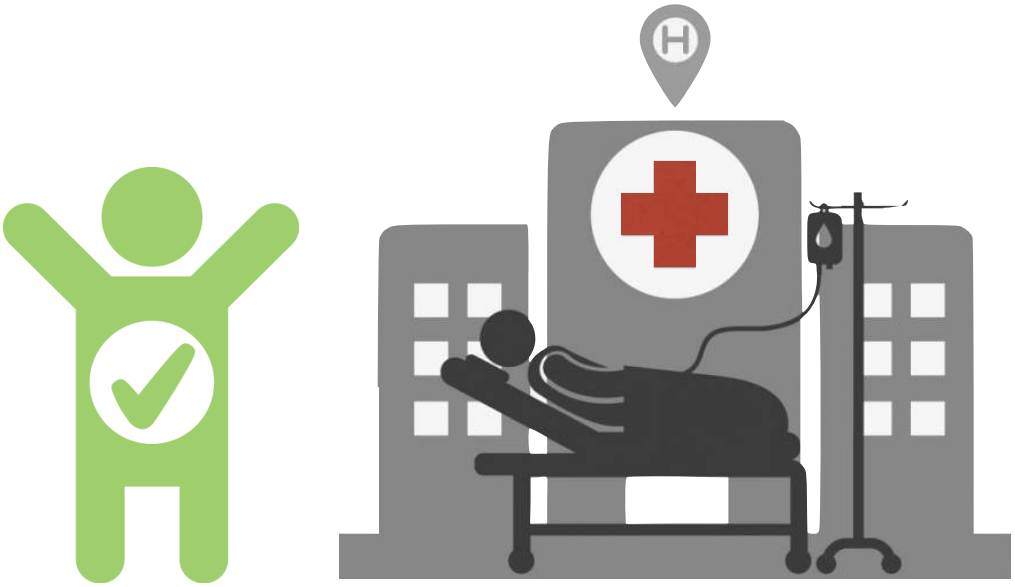


Complex Traits & Disorder Testing

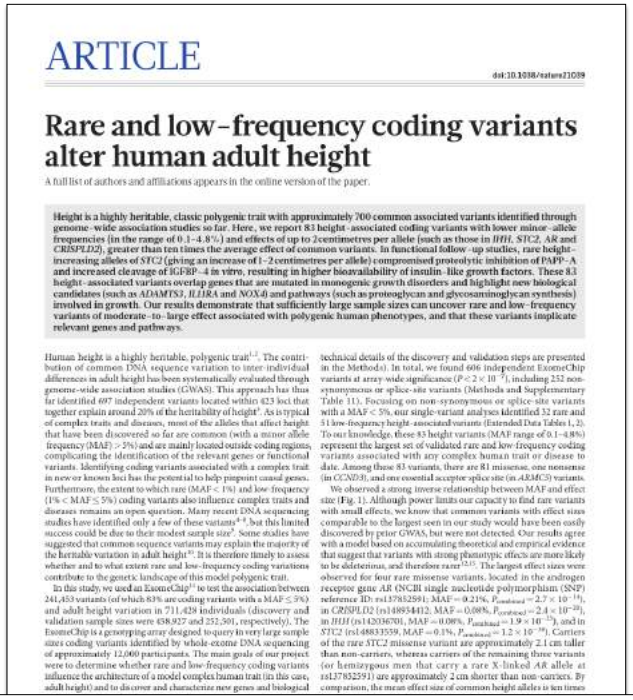
The challenge of complex traits



The challenge of complex traits



The challenge of complex traits



ARTICLE

doi:10.1038/nature21039

Rare and low-frequency coding variants alter human adult height

A full list of authors and affiliations appears in the online version of the paper.

Height is a highly heritable, classic polygenic trait with approximately 700 common associated variants identified through genome-wide association studies so far. Here, we report 83 height-associated coding variants with lower minor-allele frequencies (in the range of 0.1–4.8%) and effects of up to 2 centimetres per allele (such as those in *HHIP*, *STC2*, *AR* and *CRISPLD2*), greater than ten times the average effect of common variants. In functional follow-up studies, rare height-increasing alleles of *STC2* (giving an increase of 1–2 centimetres per allele) compromised proteolytic inhibition of PAPP-A and increased cleavage of IGFBP-4 *in vitro*, resulting in higher bioavailability of insulin-like growth factors. These 83 height-associated variants overlap genes that are mutated in monogenic growth disorders and highlight new biological candidates (such as *ADAMTS3*, *IL11RA* and *NOX4*) and pathways (such as proteoglycan and glycosaminoglycan synthesis) involved in growth. Our results demonstrate that sufficiently large sample sizes can uncover rare and low-frequency variants of moderate-to-large effect associated with polygenic human phenotypes, and that these variants implicate relevant genes and pathways.

Human height is a highly heritable, polygenic trait[1–3]. The contribution of common DNA sequence variation to inter-individual differences in adult height has been systematically evaluated through genome-wide association studies (GWAS). This approach has thus far identified 697 independent variants located within 423 loci that together explain around 20% of the heritability of height[4]. As is typical of complex traits and diseases, most of the alleles that affect height that have been discovered so far are common (with a minor allele frequency (MAF) > 5%) and are mainly located outside coding regions, complicating the identification of the relevant genes or functional variants. Identifying coding variants associated with a complex trait in new or known loci has the potential to help pinpoint causal genes. Furthermore, the extent to which rare (MAF < 1%) and low-frequency (1% < MAF ≤ 5%) coding variants also influence complex traits and diseases remains an open question. Many recent DNA sequencing studies have identified only a few of these variants[4–8], but this limited success could be due to their modest sample size[3]. Some studies have suggested that common sequence variants may explain the majority of the heritable variation in adult height[9]. It is therefore timely to assess whether and to what extent rare and low-frequency coding variations contribute to the genetic landscape of this model polygenic trait.

In this study, we used an ExomeChip[11] to test the association between 241,453 variants (of which 83% are coding variants with a MAF ≤ 5%) and adult height variation in 711,428 individuals (discovery and validation sample sizes were 458,927 and 252,501, respectively). The ExomeChip is a genotyping array designed to query in very large sample sizes coding variants identified by whole-exome DNA sequencing of approximately 12,000 participants. The main goals of our project were to determine whether rare and low-frequency coding variants influence the architecture of a model complex human trait (in this case, adult height) and to discover and characterize new genes and biological technical details of the discovery and validation steps are presented in the Methods). In total, we found 606 independent ExomeChip variants at array-wide significance ($P < 2 \times 10^{-7}$), including 252 non-synonymous or splice-site variants (Methods and Supplementary Table 11). Focusing on non-synonymous or splice-site variants with a MAF < 5%, our single-variant analyses identified 32 rare and 51 low-frequency height-associated variants (Extended Data Tables 1, 2). To our knowledge, these 83 height variants (MAF range of 0.1–4.8%) represent the largest set of validated rare and low-frequency coding variants associated with any complex human trait or disease to date. Among these 83 variants, there are 81 missense, one nonsense (in *CCND3*), and one essential acceptor splice site (in *ARMC5*) variants.

We observed a strong inverse relationship between MAF and effect size (Fig. 1). Although power limits our capacity to find rare variants with small effects, we know that common variants with effect sizes comparable to the largest seen in our study would have been easily discovered by prior GWAS, but were not detected. Our results agree with a model based on accumulating theoretical and empirical evidence that suggest that variants with strong phenotypic effects are more likely to be deleterious, and therefore rarer[12,13]. The largest effect sizes were observed for four rare missense variants, located in the androgen receptor gene *AR* (NCBI single nucleotide polymorphism (SNP) reference ID: rs137852591; MAF = 0.21%, $P_{combined} = 2.7 \times 10^{-13}$), in *CRISPLD2* (rs148934412; MAF = 0.08%, $P_{combined} = 2.4 \times 10^{-22}$), in *HHIP* (rs142036701, MAF = 0.08%, $P_{combined} = 1.9 \times 10^{-10}$), and in *STC2* (rs148833559, MAF = 0.1%, $P_{combined} = 1.2 \times 10^{-10}$). Carriers of the rare *STC2* missense variant are approximately 2.1 cm taller than non-carriers, whereas carriers of the remaining three variants (or hemizygous men that carry a rare X-linked *AR* allele at rs137852591) are approximately 2 cm shorter than non-carriers. By comparison, the mean effect size of common height alleles is ten times

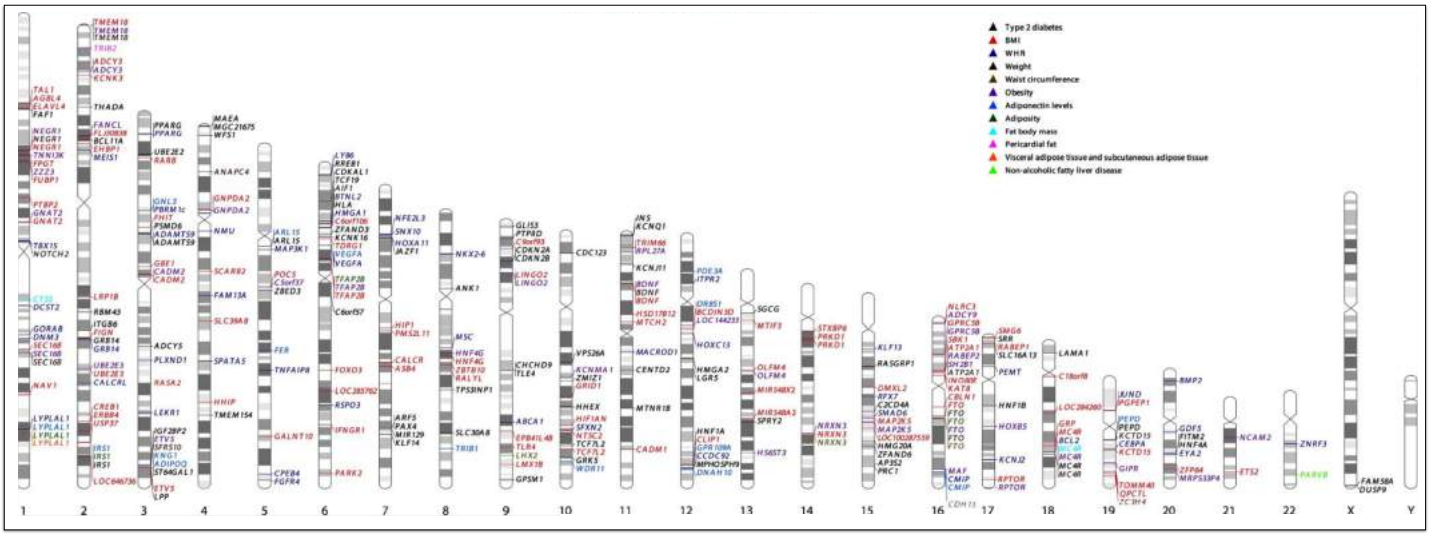
700+ genetic contributors to human height

most contribute less than 1 millimeter, but a small number of may impact up to 1 inch

nutrition also a key influence on height

The challenge of complex traits

facing a similar situations for complex disorders like type 2 diabetes



DNA regions associated with T2D and obesity-related traits

Karaderi, Drong and Lindgren, Curr Diab Rep; 15:83 (2015)

Sequencing your exome/genome

Veritas Genetics

Commercially available large-scale sequencing opportunities for consumers

Pathway Genomics

Genos

Insight genome

Explore Your DNA
Impact Your Life

HudsonAlpha
Institute for Biotechnology

Sequencing your exome/genome

Insight genome
Explore Your DNA
Impact Your Life

Clinical Component

- Initial consultation at the Smith Family Clinic
- Clinical genome analysis, linked to the clinical consultation
  - primary findings
  - secondary findings
  - pharmacogenomic results
- Follow-up clinical visit

Sequencing your exome/genome

Insight genome
Explore Your DNA
Impact Your Life

Research Component

- Invitation to participate in a HudsonAlpha research initiative to gain additional insight into the genome

Genome Guide

- **Introduction** to your genome
- **Disease chapters** - Rheumatoid Arthritis, Coronary Artery Disease, Macular Degeneration
- DNA variants associated with **other diseases**
- DNA variants linked to **non-disease traits**

Sequencing your exome/genome

Insight genome
Explore Your DNA
Impact Your Life

Research Component

- Invitation to participate in a HudsonAlpha research initiative to gain additional insight into the genome

Key questions to explore

- **How is this type of information best shared?**
- **Is this information helpful to participants and their physicians?**
- **What lessons can be learned about incorporating genomic information into routine patient care?**

issues associated with opportunistic/elective genome sequencing

- attitudes and expectations of healthcare providers and patients/ participants
  - limitations
  - applications
  - importance of education
- breadth of underlying datasets of genetic variants used for interpretation
- research/clinical boundaries are blurring, but they utilize different ethical frameworks

issues associated with opportunistic/elective genome sequencing

- meaningful consent - balancing personal autonomy and healthcare professional's duty of care
- How far should patient/participant choice guide the disclosure of clinical findings?
- What about services that do not involve a physician?
- Who has access to the data?

Next Week

Feb 7 what's in a genome?
paternity, forensic & ancestry

Feb 14 types of genetic tests, carrier & diagnostic

Feb 21 pharmacogenomics, cancer, complex disease &

Feb 28 variant validity & relevance
when to test, what lies ahead