Bone-crunching: Skeletal traits, genetic data, and their processing by supercomputer

Bailey Harrington, Bill Hill, James F Wilson, Pau Navarro, and Chris Haley
Email: bailey.harrington@igmm.ed.ac.uk
MRC Human Genetics Unit, Institute of Genetics & Molecular Medicine, The University of Edinburgh
www.ed.ac.uk/mrc-human-genetics-unit www.ed.ac.uk/igmm

Introduction

Complex Traits

- **Complex traits** are influenced by many variants across the genome, and are also affected by environmental factors.
- **Morphological traits** are complex traits to do with body size and shape.
  - Height is a common exemplar in humans.
  - Known GWAS variants do not explain all of the genetic variance component of height.
  - Dividing height into its constituent parts may help identify **novel height-associated variants** and better understand the **components of growth**.

Long Bones

- **Long bones**, specifically those of the leg, are the largest contributors to height in humans.
- Some variation in limb length exists **independent** of variation in height.

UK Biobank DXA Cohort

Cohort details

- 15,000 individuals from the UK
- Approximately 12,000 white, British individuals
- 25 million HRC-imputed SNPs

Long bone lengths were derived from dual-energy x-ray-absorptiometry (DXA) images using image registration.

This work was carried out on the Eddie supercomputer at the University of Edinburgh.

Image taken from the Lunar iDXA brochure.

Aims and Hypotheses

There are three aims to my work:

- To find variants that affect long bone length **independent of height.**
- To understand the role **rare variants** play in long bone length.
- To understand genetic factors controlling **relative proportions**:
  - of the limbs to the body; and
  - of the proximal and distal parts of the limbs.

There are several hypotheses:

- Long bone lengths also have some variation that is **independent of height**.
- After adjustment for height, different variants affect the residual variation in **arms** and in **legs.**
- Variants affecting residual variation in **individual long bones** are largely non-overlapping.
- SNPs that affect height indirectly through a direct effect on long bone length, will have **larger relative effect sizes** on long bones.
- Focus on more proximal phenotypes may result in **greater power** to find SNP-trait associations.

Heritability

REML results show long bone-length phenotypes have **high SNP-based heritability** in UK Biobank.

	h^2 (%)	se (%)	N
Height	73.58	4.31	11,539
Femur	52.68	4.57	11,160
Fibula	43.05	4.7	10,821
Humerus	54.81	4.62	10,895
Radius	33.44	4.94	10,333
Tibia	58.02	4.51	11,197
Ulna	40.25	4.97	10,309

Genome-Wide Association Analyses

- GWAS for **long bone lengths, torso length, shoulder and hip widths** have been run in the first release of the UK Biobank DXA cohort.
  - A **mixed model** was used to account for **relatedness**.
  - Height, age, sex, 10 genetic principal component, and genotyping array were included as **covariates**.
- **16 novel genome-wide significant associations** were found between individual long bone lengths or hip width and several variants.
- GWAS were also run in **separate male and female sub-cohorts**.

Trait	rsid	MAF	ß	se	p-value	Nearest Gene	Variant Type
Femur	rs12538548	0.274	1.11	0.15	4.38E-13	*IGFBP3*	intergenic
Femur	rs35278771	0.217	-0.98	0.17	3.58E-09	*CCDC91*	intron
Fibula	rs7543136	0.283	0.98	0.16	1.19E-09	*WNT4*	upstream
Fibula	rs2995060	0.041	2.12	0.37	1.11E-08	*CREG1*	downstream
Fibula	rs11720467	0.351	-0.89	0.15	7.14E-09	*MECOM*	intron
Fibula	rs4800451	0.260	0.98	0.17	5.38E-09	*CABLES1*	intron
Humerus	rs10843115	0.259	1.03	0.16	8.29E-11	*CCDC91*	intron
Humerus	rs6113414	0.427	-1.07	0.14	3.4E-14	*PAX1*	intergenic
Tibia	rs7543136	0.281	0.89	0.16	7.74E-09	*WNT4*	upstream
Tibia	rs1428217	0.427	0.79	0.14	3.03E-08	*lincRNA*	intron
Tibia	rs7105783	0.394	0.81	0.14	2.53E-08	*TEAD1*	intron
Tibia	rs4800451	0.261	0.109	0.16	1.16E-11	*CABLES1*	intron
Hip Width	rs451643	0.295	0.74	0.14	4.37E-08	*FGFR4*	upstream
Hip Width	rs6478243	0.394	0.71	0.13	2.9E-08	*ASTN2*	intron
Hip Width	rs2025009	0.375	-0.78	0.13	1.14E-09	*RADS1B*	intron
Hip Width	rs753126846	0.406	0.74	0.12	2.22E-09	*MAF*	intergenic

Sex-specific results

- **rs143384** has been identified as an **associated variant for hip width**, both under models including and not including height as a covariate.
- This **association was seen** in the **complete UK Biobank DXA cohort**, and also in the **female-specific analyses**, but not the male-specific analyses.
- This variant has **previously been associated with various hip phenotypes**, but further investigation is warranted to understand the link to sex.

Acknowledgements

- This research has been conducted using the UK Biobank Resource under Application Number 19655.
- Lucija Klaric, Andy Bretherick, and above all, Carmen Amador for answering my questions; and Paul Timmers for his software debugging assistance.
- **Financial disclosures:** My funding comes from a Principal's Career Development Scholarship through the College of Medicine and Veterinary Medicine.