**Paper meetings**

3/29/2022

* Figure 7, put together in nice layout, remove empty space, more compact
* Calcium look like outlier – more variants pass through in QC for NFE – larger distribution of z score
* Change to 90% confidence intervals – more bootstrap samples, from 100 to 1000 z-scores. Drop bottom and top 5%
* Methods section, together with derivation; put it together in the manuscript, then send to Arbel – top priority
* Matt is looking into biological info on sexual selection – down the hall
* ~~Put the new gen env figure in – change to +/-3 on each side~~
  + ~~Can change pgs one to +/- 3se on each side~~
  + ~~Keep 2 sample ttest for the pgs one~~
* ~~Change to genetic effect v testosterone level correlation (right side) to 90% confidence interval~~

4/5/2022

* Edits to selection figure done
* Do not have posterior estimate summary statistics at the moment
  + How should we store to them view? Very compressed zip files?
* Gen env pvalues: difference of ratio from 1:1
  + Z score = (ratio – 1) / ratio\_se
  + Ratio se from previous equation: (x/y)\*sqrt((x.se^2/x^2)+(y.se^2/y^2))
* Need z score table for supplement
* Choice of taking random SNPs per LD block is valid, but there are other methods that can be more powerful
  + Can use most common variants – if have higher af, have higher LD across the board
    - Can have small section: discuss in way that shows we thought about it
  + Our statements currently are about LD blocks only right now
* High variance, but might get high error for calcium; calcium has low polygenicity
* Fst looking at differences of m-f allele frequency
  + If we find variant where there is a difference
  + Variant affects longevity differently in males and females (selection)
  + Selection can be on other associated traits, trait has to affect longevity
    - Like diabetes
* Supplementary text files—corral, put in corral BOX
  + Need individual files for each figure
  + Make github available through website – make repository public
* Pvalue
  + No less or more accurate, easier to read
  + Color datapoints by whether they are significant
  + Detail method we are doing in methods
  + Don’t know if its normally distributed
  + Can look up distribution of ratio
  + Assume: Distance from 1:1 is normally distributed
  + Can make a diagram with 90% confidence interval

Supp tables needed

* Sex-specific gwas
* Posterior summary statistics
* Sex selection – filter gwas by pvalue
* Sex selection – results for other ancestry groups

4/12/2022

Matt

* Not much change in estimated A values even when remove mismatched sites
  + Each trait only had ~0-10 site removed
* Get final results from Matt – updated analysis, methods
* Papers (Kasimatis) finding sex-antagonistic selection based on M-F allele frequencies due to artifact
  + Getting mis-mapped from sex-chromosomes, high correlation between autosomes and sex chromosomes
  + Creates bias from Y chromosome (sex chromo) 🡪 mean diff in allele frequencies
  + Take out regions with sequence that maps to sex-chromosomes
  + Use BLAST
* Found overall age distribution for gnomAD – need across ancestries
* Have Fst and x-axis (gxsex genetic variance) which are both symmetric
  + Currently know for bmi, whole fat, the drive is opposite, but don’t know exact direction
  + Genetic variation selected, not focal traits; does not necessarily have to be sexually antagonistic selection
  + Simulation study – under model with different survivorship, don’t necessarily have sexually antagonistic selection, what would it look like if we performed our selection estimation
* Double check alpha – shouldn’t have female or male specific
* Testosterone mendelian randomization
  + When have actual measure of testosterone on x-axis, can’t tell for sure if it’s the cause for the genetic effect of the trait to change
  + A 🡪 B, wonder if there is causal relationship, look at correlation first
  + Mendelian randomization
    - May have confounder C that affects both A and C
    - Look at another variable D that you has causal relationship with A, therefore, the only relationship with B is if it is mediated by A, and then C will be random
    - D in this case if PGS for A bc it has causal relationship with A phenotype values \
    - PGS D may be confounded with PGS-traits B due to pleiotropy
  + Bin individuals by PGS
  + Expect to weaken relationships, can have a supplement or main
* Check correlation with age
* ~~Check how much space taking up on corral~~
* ~~T PGS, sex specific SD on x-axis or diff axes~~
* M-F effect estimate comparison
  + Split by MAF
  + Overall figure
  + LD blocks

4/19/2022

* May 6 talk at lab meeting: have a week to prepare, address her comments, be able to defend work
* Need to work on poster
* Keep prioritize manuscript
* Too long
  + ~~Move PGS to supplement, figure not as interesting; replace with small paragraph in main text~~
* Kasimatis: sex-GWAS, see different hits, say due to mismapping
  + Just can exclude those sites
* Piratsu: say different hits due to recruitment bias by sex
  + If there is GxSex interaction on participation, can affect all of our results
  + Potential issue, but most likely very small effect
  + They found weak signals for recruitment bias
* Benonisdottir: don’t find significant sex specific genetics effects based on recruitment
* Expand discussion on recruitment bias in introduction
  + Read those papers, familiarize
* M-f effect plots
  + Smaller R^2 due to noise,
  + Mash reduce SNPs from noise to null
  + Check phenotypic variance, if male biased, slope will increase vice versa; check if going in direction of variance
  + Do overall plot
* ~~Change urate to bmi or weight~~
* ~~Figure 3: log scale y axis so that the male:female amplification are even~~
* Age as confounder – why is it weird with arm fat-free mass
  + Could try version of main figure corrected for age
  + Genetic effect ~ testosterone level + age used to get correlation R^2
  + Instead use est(y\_i) = y\_i – hat(y\_i)
    - Regress y\_i on mean age across bins
    - Get correlation of (est(y\_i), testosterone level)
* Bernabeau: Gwas on one site,
  + X: AA, AT, TT
  + Y = ex. Weight sex adjusted
* **Test idea of alpha (amplification)**
  + X = genotype, polygenic score in one sex;; male score on top plot, female score on bottom plot;; expect to see similar results
    - Bin individuals by PGS
  + Y = phenotype, adjusted for sex and can also test not adjusted for sex
* PGS – prediction
  + If multiplied by same constant across big part of genotype in males, prediction is gonna be same, but just better in males
* More comments to come 😊

4/20/2022

* Comments from Molly

version of equal amplification plot only with traits that show evidence for amplification, or instead have different symbols for traits with evidence and ones without.

* + ~~Change fig 4 y axis to “effect of 1 pgs sd on phenotype”~~

4/22/2022

* Scatterplot for table 1 to compare effect sizes?
* ~~A plot of phenotypic variance ratio vs. the extent of difference in trait mean (maybe as a difference, maybe ratio, not sure). Same visual language as Fig. 3~~

4/26/2022

1. Testosterone as underlier – adjusted for age
2. Pheno variance by sex-specific PGS – figure for paper in progress
3. Pheno variance ratio by pheno mean ratio
4. M-F effect comparison
5. New figure 4 (log scale) and figure 6 (selection)

TODO:

* Perhaps table 2 move to supplement, c~~hange color for motivation section~~
* Get draft to arbel by tomorrow – can have smaller number of changes
* Units of SE for pheno variance by sex-specific PGS
* Not standardizing phenotype – clearly explain in text
  + Ex. In testosterone units, female heritability smaller than both-sex and males because male samples set the tone for testosterone
* Perhaps ~~move figure 1 to amplification section; remove reference to it in single loci section~~
* ORDER: Genetic correlation; polygenic v phenotype; mash; phenotypic distribution
  + Can leave referencing corrections for later
* Figure 7: filtered out mismapped SNPs, compare with Kasimatis for insanity check
  + Matt using read length sequencies, filter based on 90% frequence identity; move conservative
  + Cannot give convincing answer yet about recruitment bias, effect of evidence of sexually antagonistic selection -- false positive differences in allele frequencies
    - Sex differences in mortality (ex. Diabetes) recruitment bias; anything comorbid with body mass related traits
    - Body mass trait can affect tendency to go to doctor, can affect males and females differently
    - Both plausible. Esp since our sample sizes not that big
    - Replicate this in more samples (population)– can help make case more convincing case
      * Or try to replicate in samples not based on medical data ex. UK biobank
      * But trends don’t often happen across populations
    - Mellow down claims on selection in the wrong way – like hypothesis
* Testosterone as underlier adj for age – want to remove effect of age from polygenic effect; see if T levels still have relationship with polygenic affect
  + In each bin, regress phenotype ~ pgs; get beta\_pgs
  + Across bins, beta\_pgs ~ age; get beta\_age
    - Age is mean for bin
  + Estimate beta = residuals from above; get polygenic effect residualized for age
  + As long as effect of age is real, will kill effect of T
  + Can try more bins? See if signal is still killed
  + Include age analysis as is (like mendelian randomization)
    - Not robust to controls we tried
    - Could be due to lack of power – number of bins (sample size), small heritability for testosterone
    - ~~Write out in methods – where we have the mendelian randomization part~~
* Show if results are affecting mean or variance differences in males/females
  + Careful to make claims, because can be affecting each other
* ~~Send M-F effect comparison plot over slack~~
* Trait value v sex-specific polygenic score summary across traits – overall figure
  + Two versions of figure – one for males, one for females
  + Can possibly replace the examples
  + Explain in caption that R^2 are calculated based on raw data
  + In the new figure let’s have 2 rows (sex-specific PGS) by 4 columns (traits) for: Albumin (no systematic amplification), height (because model trait), blood pressure , hip circumference.
* Poster and presentation talk few days from now
* Fix supplements

Selection

* Get table of number of loci removed – between 10% – 20% of site removed
  + Kasimatis pretty conservative – we also prob filtered out in QC step
* Table: Replication of mapping to sex chromosomes on genotyping array sites
  + Identified by kasimatis vs identified by us as mismapping: Yes/No 2x2 table
    - Numbers and percentages in table
* Keep version of just genotyping array sites as supplement
* Write out new methods and supplement
* Current one is 90% CI

4/29/2022

Poster:

* Prepare elevator pitch ~1.5 min – motivation -> results
  + Streamlined like bernabeau poster
* To prepare for grilling – can add a few more plots to help explain outside of the main story line
* Necessary – title, affiliations, emails, QR code for bioRchive preprint
* Can print out here or print there – paper
* Slack – poster for online participants

Talk next Friday:

* Motivation, more on mash methods, less on testosterone underlier and PGS, replace with selection
  + Focus on methods specifically behind mash – really understand that
  + Tie to math
* Selection part: study and go over with Arbel next week
* Paper: add equation relating modeling effect size from mixture in mash figure
* May have pointed questions – more about discuss than defend
  + Ask them to clarify questions
* 1 hr meeting – make presentation short

5/2/2022

* Do you have a sex-specific intercept when you analyze the data?
* ~~Can we include a supp table with the actual values? (heritability and correlation)~~

5/3/2022

* Large role of GxSex? Associations from GWAS based on existing genetic variation (common) – not strongly selected, vs rare fixed differences
  + Not all polymorphic sites that have diff effects in males and females
  + Is GxSex have an important role, has role in other species; important motivation, but not answered
* Assertion rather than methods introduction in poster
  + Highlight evidence for amplification part
* QR code

5/5/2022

Poster

* Just genetic correlation / or imperfect genetic correlation instead of negative (since negative is obvious)
* Male and female labels to people figure – perhaps physique?
* Male and female specific error in the GWAS equations
  + Use the Betas for the covariance matrix in place of M, F
* Add selection part?
  + Model and some zscores across traits (cut of it)
* Systematic
* Nearly similar distribution of autosomal alleles
* My contact details on poster
* Twitter consider

IGF-1 and testosterone

<https://journals.physiology.org/doi/full/10.1152/japplphysiol.00599.2016>

https://journals.lww.com/nsca-jscr/fulltext/2020/05000/sex\_differences\_in\_resistance\_training\_\_a.30.aspx

* Most studies show significant elevations in recovery after RE of testosterone in men, and no significant acute elevations in women
  + T level gains more temporary in women, significant decrease in comparison to mean

<https://academic.oup.com/jcem/article/90/5/2941/2836965>

* T increased and estradiol decreased total IGF1 concentrations
* Higher free IGF-1 concentrations in men than women receiving exogenous testosterone steroids
  + Have higher peak IGF-1 concentrations
* Infusion of rhOGF-1 induce higher IGF-1 concentrations and more rapid increases in people with more testosterone
* Sex steroids modulate bioavailability of IGF-1 in adults

https://pubmed.ncbi.nlm.nih.gov/19318219/

* Women higher insulin sensitivity than men
  + Men have higher postprandial insulin after eating than women

<https://academic.oup.com/jn/article/132/12/3799S/4712099>

* Estrogen, through ER, induce expression of IGF1
* Binding of IGF-1 to its receptor activated estrogen receptor

<https://www.frontiersin.org/articles/10.3389/fcell.2021.630503/full>

* IGF-1 plays compensatory role for both ER and AR when levels of ligand are low
* Role of estrogen complicated: stimulatory effects of GH/IGF1 axis, but localized estrogen synthesis in peripheral tissue is inhibitory

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6606898/>

* FOXO3

5/17/2022

* Mendelian testosterone underlier
  + Replace y-axis label pgs with pgs sd

6/6/2022

* Tasks when paper reviews finish ~ month or two; can submit only one at a time
* Vagheesh’s lab: MRI data – develop deep learning algorithm to do segmentation to take anatomy measurements for 40K individuals
  + Asked to take sex-specific GWAS measurements – BOLT-LMM
  + Can be put into mash as new analysis to buff up our current analysis – new raw data

7/6/2022

Reviewer 1

* Make clear in introduction while sex-stratified GWAS are not new, genome wide testing of distribution of effects finds new results
* Amplification
  + Harper et al. Evolution 2021; Karp et al. Nature Communications 2017
* Need to more clearly support conclusion, not overstate it; and address the other models in table 1 more
* Using raw v normalized data
* Explain more clearly about why not much increase in prediction

Reviewer 2

* Explain mixed models more clearly
* Explain amplification heatmap better- compare null and nonnull
* Why amplification better fit than sex-specific effects
* Estimate significance in results rather than counting
* Other study found more sex-specific effects and opposite effects, and very few diff in magnitude
  + Traglia et al PLoS Genetics 2022
  + Selected most sex-heterogenous SNPs
* Small note: using female vs male polygenic score in testing evidence for amplification
* Submit response to reviewers: inline comments, what we have done to fix it
* Follow guidelines for Cell
* I will focus on analysis – response to reviewers
* Arbel will meet with editor: explain game plan
* Graphical abstract – Cell specific

**Analysis for Carrie**

* Make idea of modeling covariance structure easier to digest for readers
* Simulation framework we already have –
  + Simulate null model and not getting amplification effects (in environmental variance supplement)
  + Also use simulation framework – not under the null
    - Have simulation where you have a lot of small effects (ex 14% effects female specific, 14% male, 76 or 86%)
    - Can also add more scenarios to match Table 1
    - Just need new method to sample mash effect sizes
    - Show specifically for R2’s scenario, would also want to show condensed matrix, equivalent of Fig 3C
* Traglia et al: take SNPs detected as sex-heterogenous, they see more conditionally neutrality or opposite effects
  + Simulation to show how top SNPs look in mash simulation
    - Get difference in distribution, ascertainment bias
  + Read Traglia paper to see how ascertainment is done to get the sex-specific top SNPs
* Polygenic score: add halved sample bars – add 4th bar to male only and female only
  + What could underlying architecture be that we don’t we a benefit in prediction – Eric
  + Could potentially include preliminary data from Eric to show that we are not hand waving
  + Do analysis for additive model with half the sample (sex-specific sample size)
    - Extra points to show results from different sample sizes, show where it intersects with prediction accuracy of mash model
* How to support amplification more: flesh out evidence for amplification
* Read Cell guidelines, see how far from requirements
* Reformat to adhere to STAR Methods