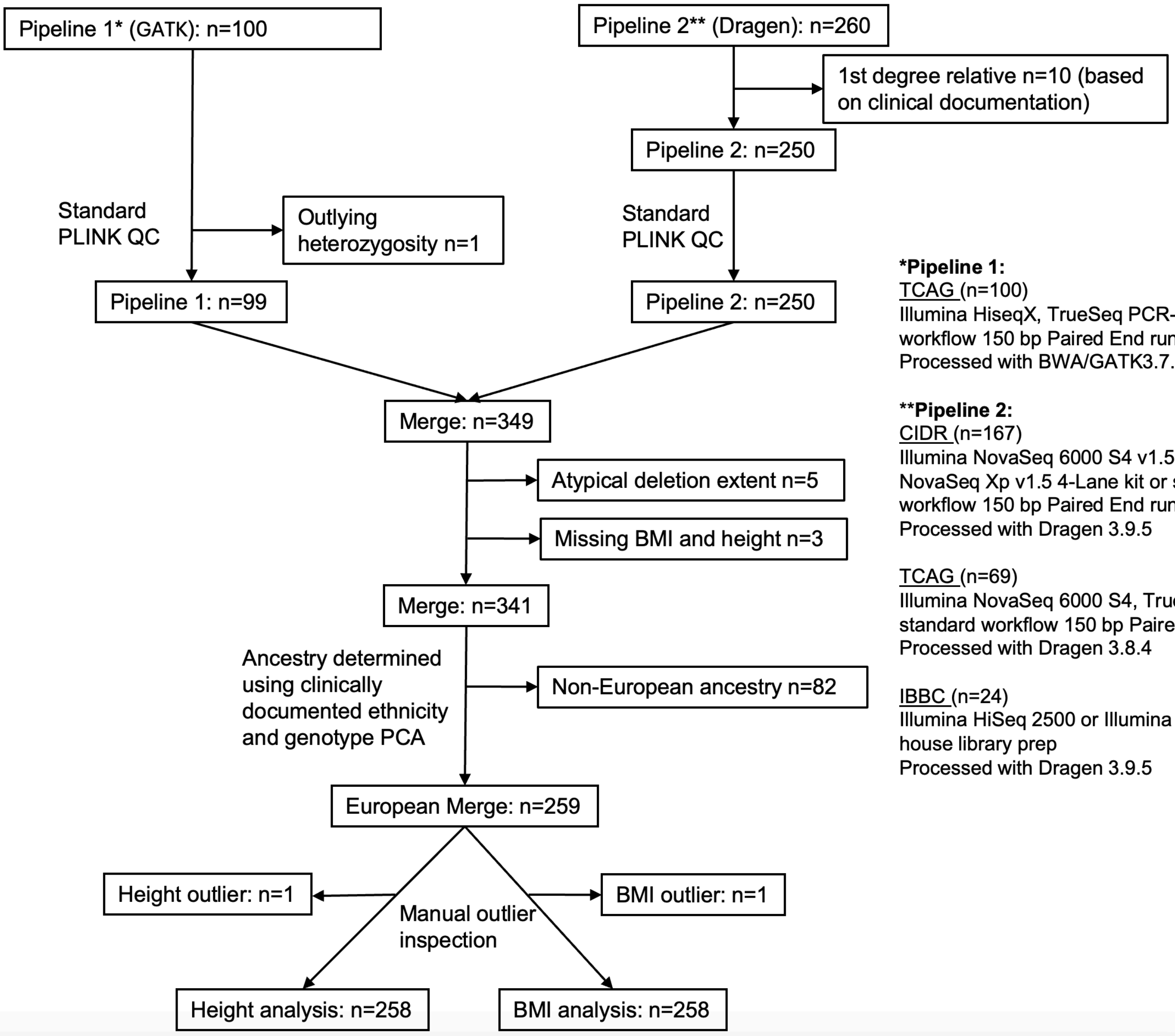
**Supplementary Table 1**. Comparison of receiver operating curve results for logistic regression models predicting short stature in adults with a 22q11.2 microdeletion.

| **Model predictor variables** | **AUC** | **Optimal threshold (Youden's)** | **Specificity** | **Sensitivity** | **p (vs AUC of covariates-only model)** | **Z (vs AUC of covariates-only model)** |
|---|---:|---:|---:|---:|---:|---:|
| Covariates-only | 0.6243 | 0.1792 | 0.591 | 0.640 | – | – |
| PRS only | 0.7314 | 0.2040 | 0.697 | 0.700 | 0.1124 | -1.5877 |
| PRS + covariates | 0.7791 | 0.1752 | 0.649 | 0.820 | 0.0011 | -3.2626 |

*AUC* area under the curve. Covariates used for the logistic regression models were sex, age, 22q11.2 microdeletion extent, moderate-severe congenital heart defect, moderate-severe intellectual disability, and psychotic illness.



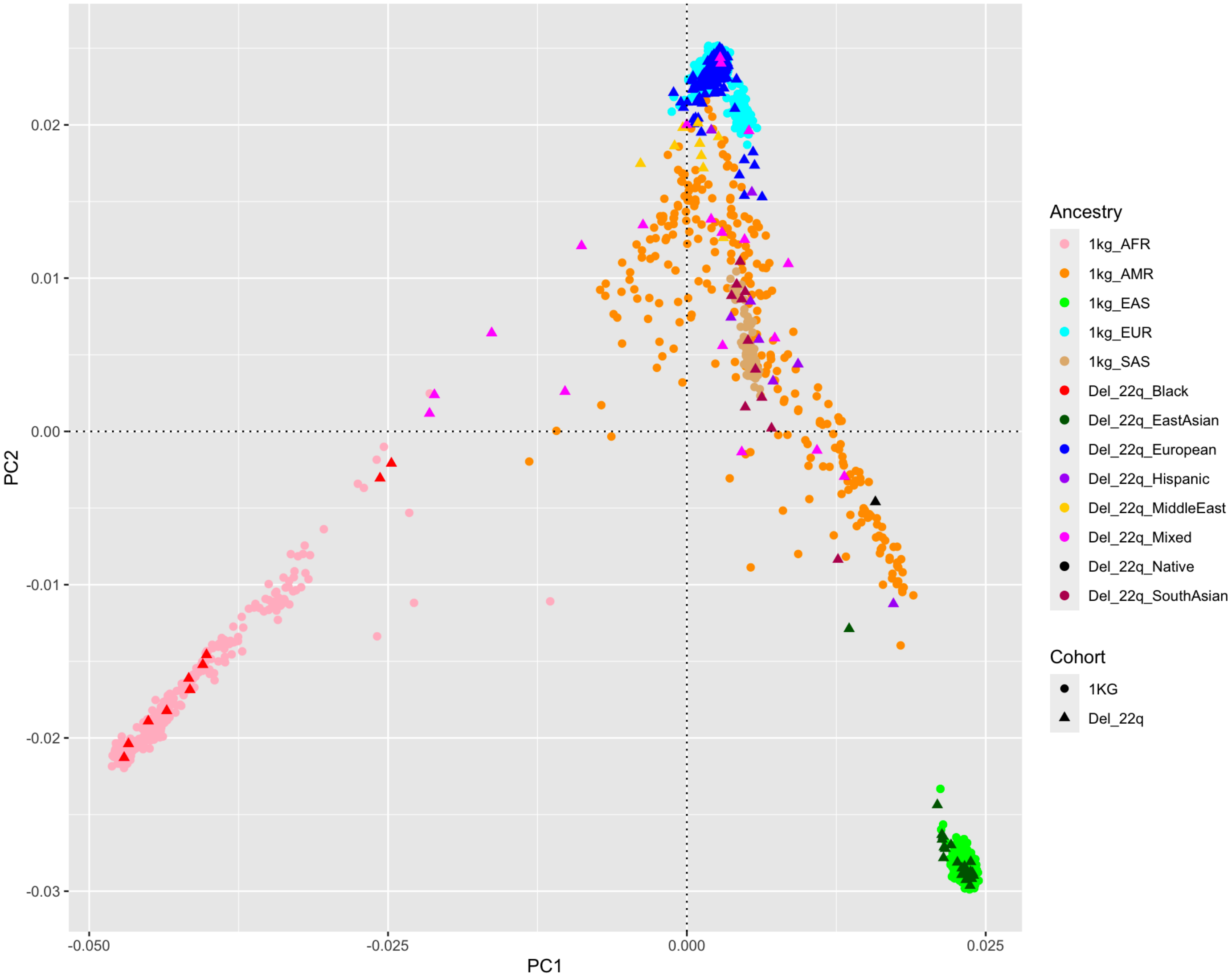
***Pipeline 1:**
TCAG (n=100)
Illumina HiseqX, TrueSeq PCR-free standard workflow 150 bp Paired End runs.
Processed with BWA/GATK3.7.

****Pipeline 2:**
CIDR (n=167)
Illumina NovaSeq 6000 S4 v1.5 Reagent Kit, NovaSeq Xp v1.5 4-Lane kit or standard workflow 150 bp Paired End runs
Processed with Dragen 3.9.5

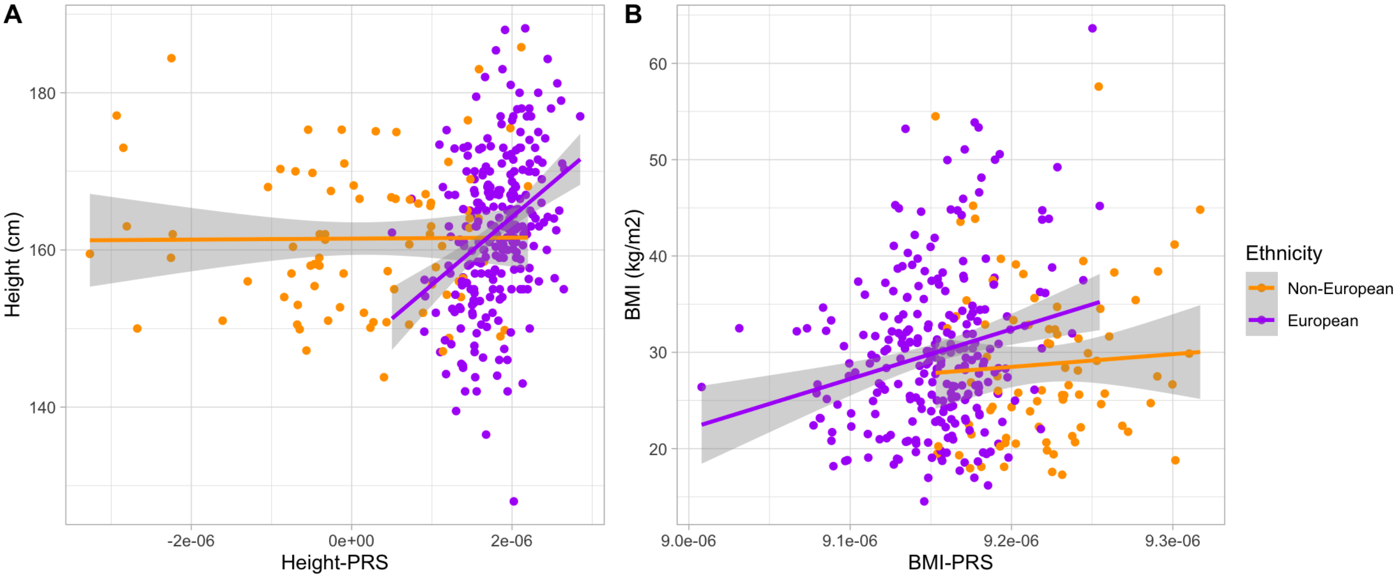
TCAG (n=69)
Illumina NovaSeq 6000 S4, TrueSeq PCR-free standard workflow 150 bp Paired End runs
Processed with Dragen 3.8.4

IBBC (n=24)
Illumina HiSeq 2500 or Illumina HiSeq X, in-house library prep
Processed with Dragen 3.9.5

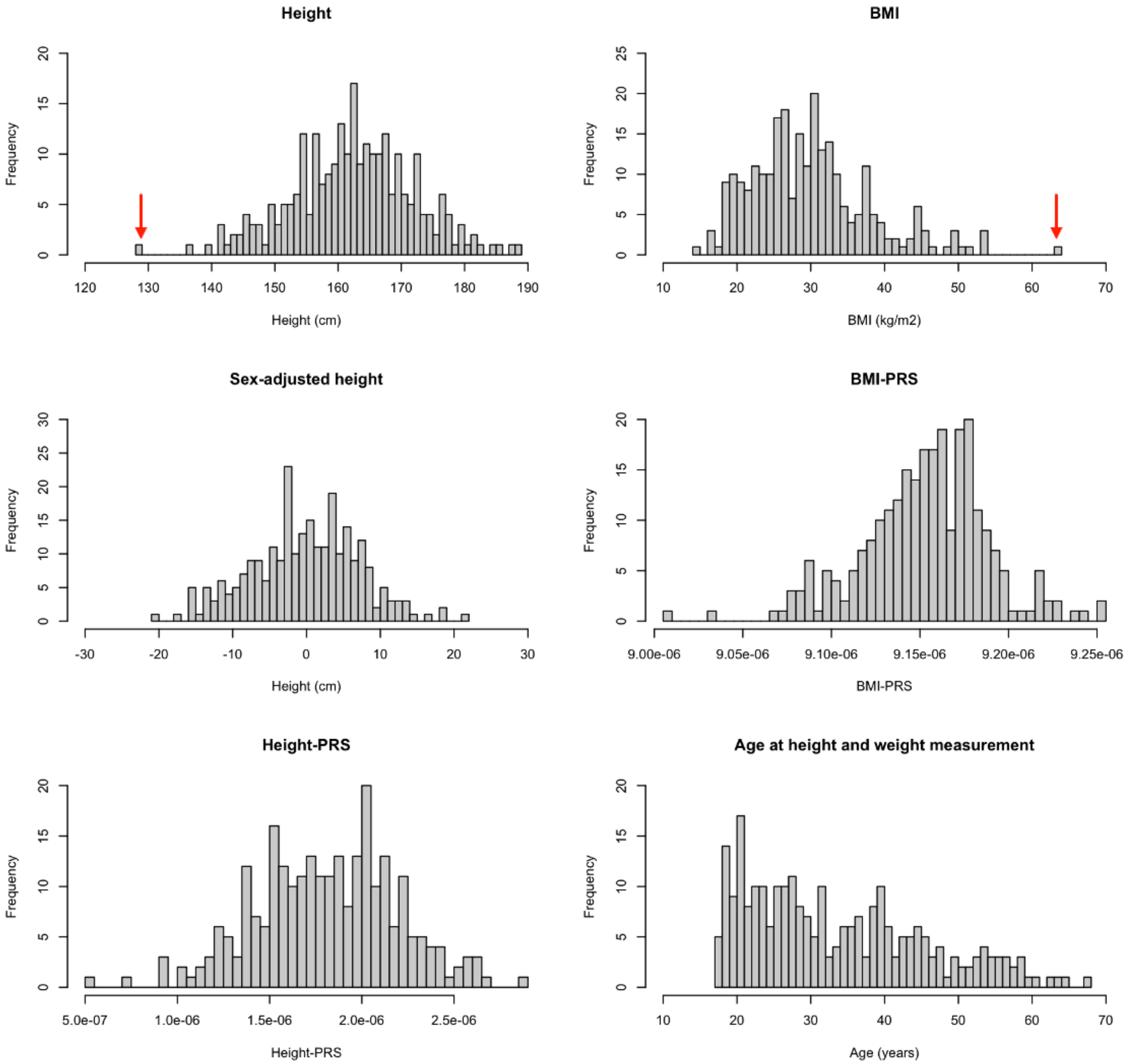
**Supplementary Figure 1**. Overview of the study design. Each *n* refers to the number of individuals with a 22q11.2 microdeletion age ≥17 years. The column on the right lists the facility each set of samples was sequenced at and a summary of the sequencing platform and variant calling method. *TCAG* The Centre for Applied Genomics (Toronto, Canada); *CIDR* Center for Inherited Disease Research (Johns Hopkins University, Baltimore, United States); IBBC International Brain and Behaviour Consortium (HudsonAlpha Genome Sequencing Center, Huntsville, United States). *PCA* principal component analysis. See Supplementary Figure 4 for individuals identified as outliers based on height or BMI.



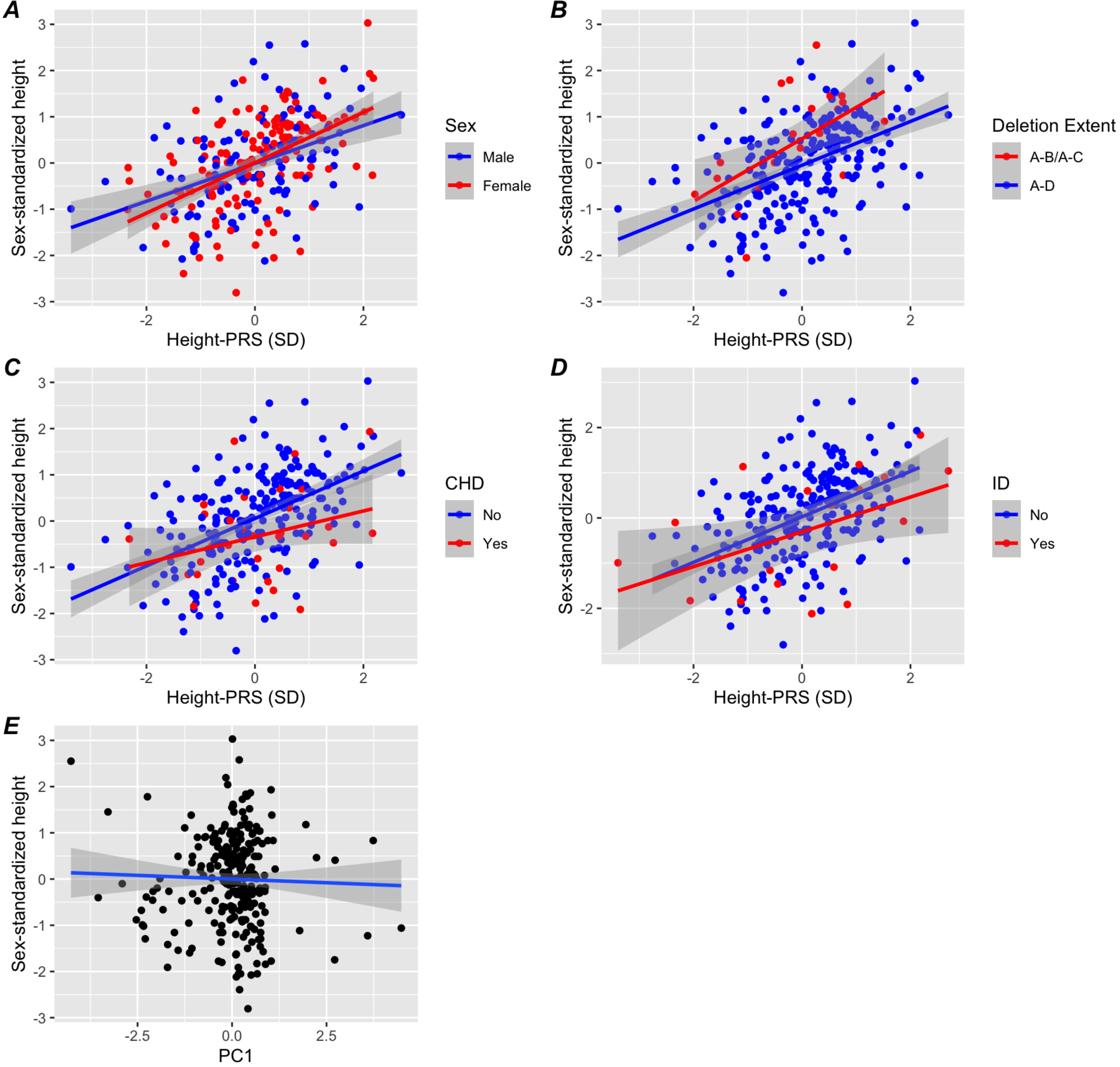
**Supplementary Figure 2.** The first two principal components of 341 individuals with a 22q11.2 microdeletion ("Del_22q" in the Cohort key and represented by triangles) and individuals who were part of a 1000 Genomes ancestry reference panel ("1KG" in the Cohort key and represented by circles) (http://www.tcag.ca/tools/1000genomes.html).



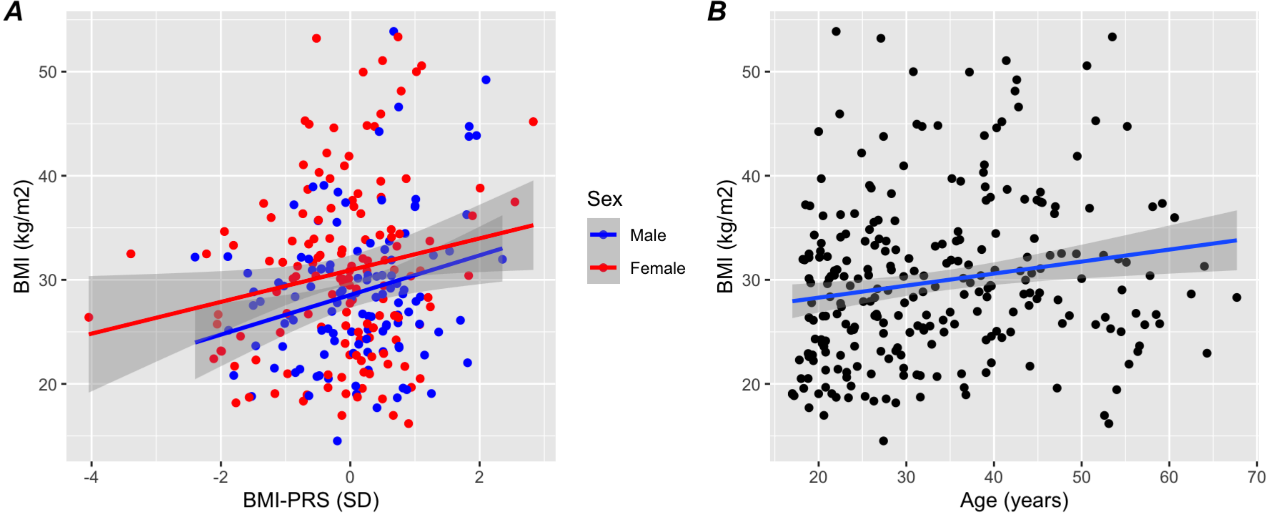
**Supplementary Figure 3.** Scatterplot and linear associations between the polygenic risk score for height (A) and body mass index (B) for 341 individuals aged ≥17 years with a typical 22q11.2 microdeletion, stratified by European vs non-European ancestry. Fitted lines were generated using linear regression within European and non-European subgroups.



**Supplementary Figure 4**. Histograms showing the distribution of height (cm) and BMI (kg/m$^2$) and the polygenic risk scores (PRS) for height and BMI in 259 individuals age ≥17 years and of European ancestry with a 22q11.2 deletion. The red arrows indicate the one individual in each plot identified as an outlier for height or BMI, that was excluded from the primary analyses of the respective trait. Note that the outlier identified for height (128 cm) had a BMI of 20.5 kg/m$^2$ and was not an outlier in the BMI distribution. Sex-adjusted height is the difference (cm) between the individual's height and the mean height of the individual's sex. The outlier for height was removed in sex-adjusted height panel only (n=258).



**Supplementary Figure 5.** Scatterplots showing the effect of sex and significant covariate predictors of height in the multivariable linear regression model (Table 2). For categorical variables, a scatterplot was generated showing the relationship between sex-standardized height and height polygenic risk score (PRS), with fitted lines generated by linear regression stratified by levels of the categorical variable (sex (A), 22q11.2 deletion extent (B), moderate-severe congenital heart defect (CHD) (C), and moderate-severe intellectual disability (ID) (D)). The first principal component of ancestry (PC1) was also plotted against height (D).



**Supplementary Figure 6.** Scatterplots showing the effect of significant covariate predictors of BMI in the multivariable linear regression model (Table 2). Fitted lines generated by linear regression were stratified by sex in a plot of BMI (kg/m$^2$) and BMI polygenic risk score (PRS) (A), and age at BMI measurement was plotted against BMI (kg/m$^2$) (B).