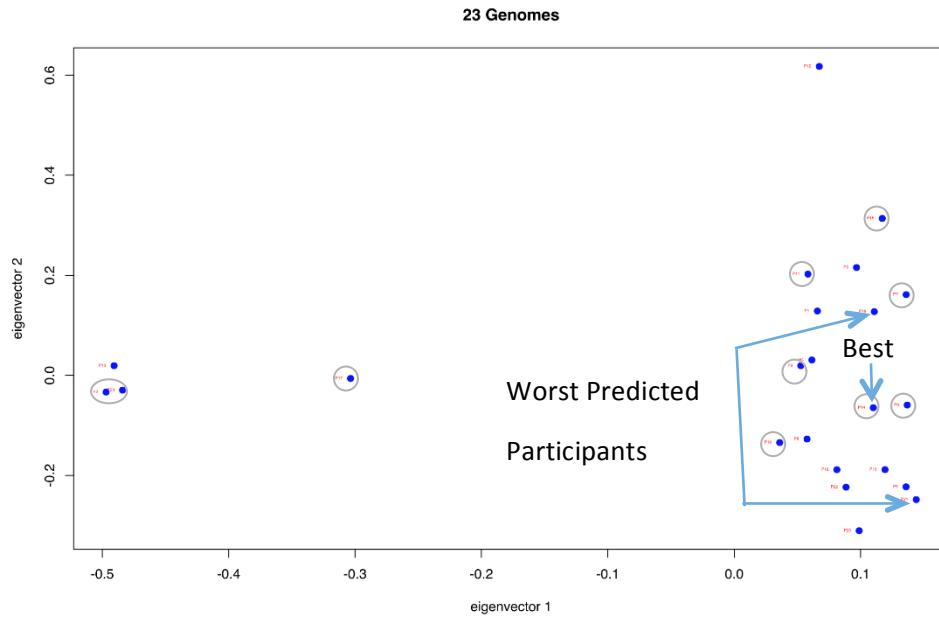
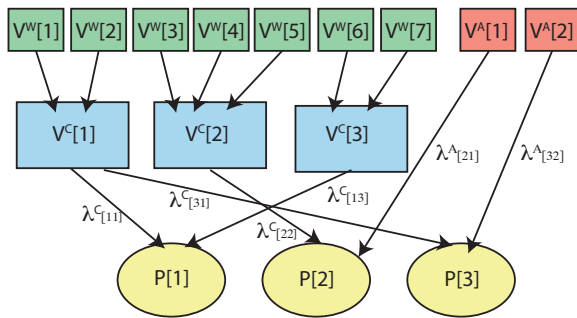


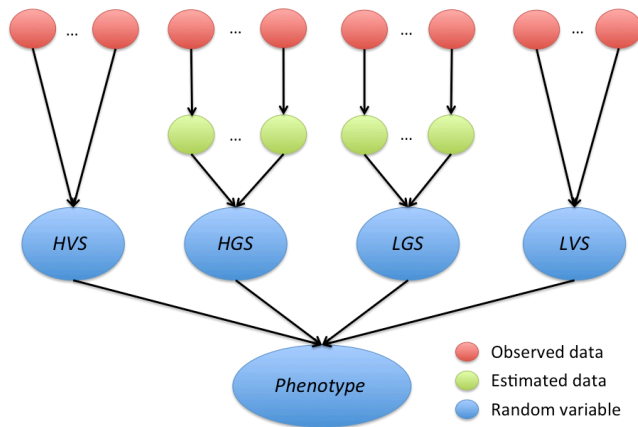
Supp. Figure S1. Eigenvectors of the 23 genomes of PGP 2015 from an analysis of ancestry.



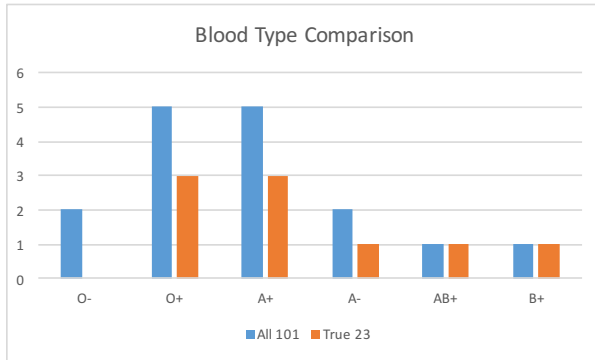
Supp. Figure S2. Toy example of Group1's predictive model. The causes of a phenotype $P[i]$ are variants $V^C[j]$ (hidden causal variants underlying GWAS hits $V^W[q]$) and variants $V^A[1]$ (variants with database annotations associated with phenotype $P[i]$). The association between variant a $V^C[j]$ or $V^A[1]$ and a phenotype $P[i]$ is represented by a directed edge, with an associated weight (penetrance) $\lambda^C[ij]$ or $\lambda^A[i1]$. $\lambda^C[ij]$ is computed by averaging the penetrances of the $V^W[q]$.



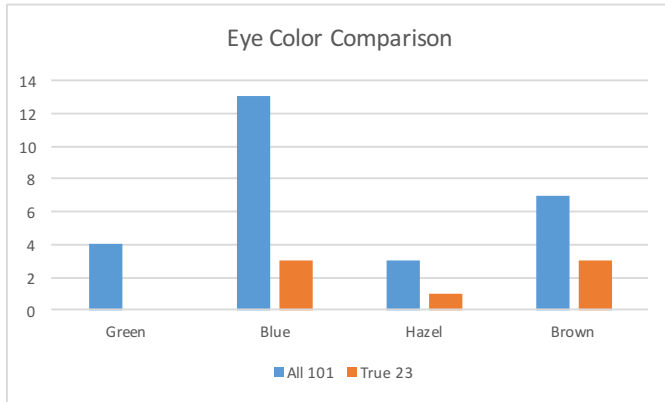
Supp. Figure S3. Bayesian network for predicting a phenotype. Red nodes represent an individual's genotype at a segregating site. Green nodes represent genes. Circular blue nodes represent unknown mechanisms that determine the penetrance of affected variants and genes on a phenotype. Red nodes are observed in sequencing data. Green nodes are computational predictions ("estimated data"). Blue nodes are random variables.



Supp. Figure S4. Participant count comparison of blood types between 101 profiles and 23 true profiles of PGP 2015.



Supp. Figure S5. Participant count comparison of eye color between 101 profiles and 23 true profiles of PGP 2015.



Supp. Table S1. Participant profile information ranked by submissions of PGP 2015.

Rank	Profile	Blood type	Left eye	Maternal grandmother	Paternal grandmother	Paternal grandfather	Maternal grandfather	Age	Gender	Race
1	14	AB +	light brown	Italy	Russian Federation	Poland	Italy	50-59	Female	White
2	12	N/A	N/A	Netherlands	US	US	Netherlands	1992	Male	White
3	13	N/A	N/A	China	China	China	China	1975	Male	Asian
4	5	A +	Blue	US	US	Ireland	US	1948	Male	White
5	6	N/A	N/A	UK	Germany	Ireland	UK	30-39	Male	White
6	11	A +	aqua	US	US	US	US	1991	Neutrois	White
7	15	O +	brown	US	US	US	US	30-39	Female	White
8	7	B +	hazel	US	Other	Germany	US	1947	Female	White
9	9	A +	blue	Germany	Germany	Germany	Germany	40-49	Male	White
10	23	N/A	N/A	Viet Nam	Viet Nam	Viet Nam	Viet Nam	1982	Male	Asian
11	17	N/A	N/A	Peru	Peru	Peru	Peru	1989	Male	Hispanic /Latino
12	3	O +	N/A	US	US	US	US	40-49	Male	White
13	4	N/A	N/A	Canada	US	US	US	1983	Male	White
14	8	N/A	N/A	US	Turkey	Turkey	US	1967	Male	White
15	18	N/A	N/A	US	France	France	US	1974	Female	White
16	22	N/A	N/A	US	US	Canada	US	21-29	Male	White
17	20	N/A	N/A	US	US	US	US	21-	Male	White

								29		
18	1	N/A	N/A	US	US	Canada	US	21-29	Male	White
19	19	N/A	N/A	US	US	US	US	1951	Male	White
20	2	N/A	N/A	China	China	China	China	1988	Male	Asian
21	10	O +	dark brown	US	Puerto Rico	Puerto Rico	US	1988	Male	Hispanic or Latino, White
22	21	N/A	N/A	US	US	Other	Other	1981	Male	American Indian/Alaska Native, White
23	16	A -	blue	US	Denmark	Denmark	Canada	50-59	Male	American Indian/Alaska Native, White

Supp. Table S2. The phenotypes that were chosen by the Group5's submission in PGP 2015.

k	Phenotype	Prediction Method	SNPs Used	Covariates
1	sex	ad hoc	presence of any SNP on chr Y	
2	Baldness	logistic regression	rs2497938	sex, age
3	Astigmatism	logistic regression	rs10496034, rs1658442, rs7534824	age
4	osteoarthritis	logistic regression	rs1417066, rs3757837	sex, age
5	blood type rh blood type ABO	logistic regression conditional multinomial	rs590787 rs8176719, rs8176747	
6	eye color blue	logistic regression	rs12913832	
7	Floaters	6 nearest neighbors	random 1000 SNPs not on chr Y	
8	high triglycerides	6 nearest neighbors	random 1000 SNPs not on chr Y	
9	Gingivitis	6 nearest neighbors	random 1000 SNPs not on chr Y	