

This document contains the list of ambiguities in the OpenSNP phenotype file, and to handle these issues; we transformed the phenotype's values into two classes.

In postTransformation, we considered four values: Class 1, Class 2, U (Unsure, means we are not sure whether people who selected this phenotype falls in Class1 or Class2), and x (means this phenotype is removed because its values contain too many instances or a high difference between the number of cases and controls).

Duplicate values

Same values for Lactose intolerance: Lactose-intolerant, lactose-intolerant, No

Same values for Lactose tolerance: Lactose-tolerant, lactose-tolerant, No

Phenotype	Lactose intolerance	Lactose intolerance
5668	-	-
255	Lactose-tolerant	Lactose-tolerant
116	Lactose tolerant	Lactose-tolerant
96	Lactose-intolerant	Lactose-intolerant
69	Partially lactose intolerant - surfaces with a lot of dairy products in one day	U
56	lactose-tolerant	Lactose-tolerant
34	FALSE	Lactose-tolerant
21	No	Lactose-tolerant
19	Genetically intolerant but drink raw milk and eat lots of dairy.	U
16	Genetically intolerant partial intolerance.	U
10	Slight lactose intolerance	Lactose-intolerant
9	lactose-intolerant	Lactose-intolerant
5	allergic to all forms of dairy	Lactose-intolerant
5	Only with cows milk products	U
4	Genetically tolerant-extrememly lactose intolerant	U
3	Lactose intolerant	Lactose-intolerant
2	Problems if i drink more 1 liter milk	Lactose-intolerant
2	Lactose intolerance appearing with age at mid 40s.	U

2	Lactose tolerant till my 40s then became very intolerant. genetics says i am intolerant.	U	
2	lactose tolerant till my 40s then became very intolerant. Genetics says I am Intolerant.	U	
2	rs4988235	U	
2	lactose tolerant	Lactose-tolerant	
2	AA	U	
1	No but I seem to have less inflammation in joints if avoid milk products	U	
1	severe GI pain	U	
1	Allergic to all forms of dairy	Lactose-intolerant	
1	Genetically tolerant but am intolerant	U	
1	Severe gi pain	U	

High difference between the number of cases/controls.

Instances of not diagnosed with otosclerosis >> Instances of Diagnosed with otosclerosis

Phenotype	Otosclerosis	x
6328	-	x
73	Not diagnosed with otosclerosis	x
3	Diagnosed with otosclerosis	x
1	Caucasian	x

Unclear answers.

People who answer are not sure whether it's boldness or baldness.

Phenotype	Boldness type
6236	-
92	I am not bald
11	Front and rear
11	Hey theres one of your bugs boldness =baldness ??? me im not bald at all
11	Average
10	Mid-frontal
8	No I'm pretty timid not bold)
7	Rear
6	No i'm pretty timid not bold)
3	Arial

3	I'm not very bold
2	front and rear
2	Bald
1	rear
1	mid-frontal
1	as bold but not as bald as an eagle

Phenotypes which can have more than 2 instances.

Multiple values for eye pigmentation, so we discarded such phenotypes.

Phenotype	Eye pigmentation	x
6289	-	x
26	Blue	x
21	Brown	x
20	Hazel dark green	x
16	Changes blue/green/grey	x
10	Green	x
5	Hazel light green	x
3	Rs16891982gg	x
2	Blue-grey	x
2	Heterochromia	x
2	changes blue/green/grey	x
1	hazel light green	x
1	DARK BLUE	x
1	eye color what do you think	x
1	Hazel/Yellow	x
1	Light Gray Blue	x
1	green brown	x
1	Hazel/yellow	x
1	Hazel green	x
1	No	x

Phenotypes for which people are unsure.

Phenotype	chronically sore neck glands	chronically sore neck glands
6359	-	-
33	No	No
7	Yes almost always	Yes
1	no	No
1	yes almost always	Yes
1	not sore but enlarged for decades	U
1	in my youth but not since moving to a warmer climate	U
1	Regular flares -seems to especially occur when having histamine response or during PMS	U
1	Not sore but enlarged for decades	U