

## GWAS

(по материалам <https://education.23andme.com//how-to-do-a-gwas/>)

For one trait, bitter taste ability, the results from last year are as follows:

X4988235

bitter	AA	AG	GG	NULL
bitter_no	2	3	7	1
bitter_yes	4	9	11	1

=====

X7495174

bitter	AA	AG	GG	NULL
bitter_no	9	2	1	1
bitter_yes	13	8	3	1

=====

X713598

bitter	CC	CG	GG	NULL
bitter_no	7	4	1	1
bitter_yes	1	13	10	1

=====

X17822931

bitter	CC	CT	NULL	TT
bitter_no	6	3	1	3
bitter_yes	13	7	1	4

=====

X4481887

bitter	AA	AG	GG	NULL
bitter_no	0	6	6	1
bitter_yes	6	10	8	1

=====

Now, let's determine the association between bitter taste and the first SNP (rs4988235) using a  $\chi^2$  test:

X4988235

bitter	AA	AG	GG	NULL
bitter_no	2	3	7	1
bitter_yes	4	9	11	1

=====

1. Get observed allele counts • There are 2 As for AA homozygotes and 1 A for AG heterozygotes. Similarly, there is 1 G for AG heterozygotes and 2 Gs for GG homozygotes.

	Bitter taster	Non-taster	TOTAL
A	$2*4+1*9 = 17$	$2*2+1*3 = 7$	24
G	$2*11+1*9 = 31$	$2*7+1*3 = 17$	48
TOTAL	48	24	72

2. Get observed frequencies • Normalize by the total

	Bitter taster	Non-taster	TOTAL
A	$17/72 = 23.6\%$	$7/72 = 9.7\%$	33.3%
G	$31/72 = 43.1\%$	$17/72 = 23.6\%$	66.7%
TOTAL	66.7%	33.3%	100%

3. Get expected frequencies • Treat events as independent and calculate the contingency table. • I.e.  $P(\text{Bitter taster} \cap A) = P(\text{Bitter taster}) * P(A)$

	Bitter taster	Non-taster	TOTAL
A	$.333 * .667 = 22.2\%$	$.333 * .333 = 11.1\%$	33.3%
G	$.667 * .667 = 44.5\%$	$.333 * .667 = 22.2\%$	66.7%
TOTAL	66.7%	33.3%	100%

4. Get expected counts • Multiply the frequency by the total allele counts

	Bitter taster	Non-taster	TOTAL
A	$.222 * 72 = 15.98$	$.111 * 72 = 8.00$	23.98
G	$.445 * 72 = 32.04$	$.222 * 72 = 15.98$	48.02
TOTAL	48.02	23.98	72

5. Use the  $\chi^2$  equation. Then lookup  $\chi^2$  value to get p-value using 1 degree of freedom (can plug this number into an online calculator).

$$\chi^2 = \sum (O - E)^2 / E$$

O = observed

E = expected

$$\chi^2 = (17-15.98)^2/15.98 + (7-8)^2/8 + (31-32.04)^2/32.04 + (17-15.98)^2/15.98 = 0.07 + .13 + .03 + .07 = 0.3$$

$$P(\chi^2=0.3) = 0.58$$

ВЫВОД: нет значимой ассоциации

## Поправка на множественное тестирование

Все снипы выше были исследованы на микрочипах с 500 000 маркерами. Если бы исследовали только один снип, то найденная ассоциация была бы значимой при уровне  $\alpha=0.05$ . Но в случае тестирования 500 000 снипов,  $p\text{-values} = 0.05/500\,000 = 10^{-7}$ . Таким образом, мы делаем вывод, что rs4988235 не ассоциируется статистически значимо с восприятием горького вкуса с поправкой на множественное тестирование.

## Точный тест Фишера

*Пример кода в R*

```
chisq.test(matrix(c(top-left, top-right, bottom-left, bottom-right), nrow=2), correct=FALSE)
```

```
•chisq.test(matrix(c(17,7,31,17), nrow=2), correct=FALSE)
```

Normally we would not set `correct=FALSE`, but on paper we weren't applying a continuity correction, so this will give us the results we calculated by hand.

```
•fisher.test(matrix(c(17,7,31,17), nrow=2))
```

*Пример кода на питоне*

```
from scipy.stats import chi2_contingency
stat, p, dof, expected =chi2_contingency([[17, 7],[31, 17]],
    correction=False)
```

```
print(p)
```

#точный тест Фишера

```
from scipy import stats
```

```
table=[[17,7],[31,17]]
```

```
oddsratio, pvalue=stats.fisher_exact(table)

print(pvalue)

print(oddsratio)
```

Или онлайн-калькулятор:

<http://medstatistic.ru/calculators/calchi.html>

### Отношение шансов (Odds ratios)

Когда мы обнаружили снип, статистически значимо ассоциированный с признаком, мы можем посчитать отношение шансов для данного снипа.

bitter	A	G
bitter_no	17	7
bitter_yes	31	17

=====

### Odds Ratio

Allelic odds ratio: ratio of the allele ratios in the cases divided by the allele ratios in the controls

Берем отношение A/G в группе не чувствующих горечи - контроль  
 $17/7=2.42$

Берем отношение A/G в группе, чувствующих горечь - случай  
 $31/17=1.82$

Odds ratio =  $1.82/2.42=0.75$

То есть люди с аллелью A в 0.75 раз менее вероятно будут ассоциированы с способностью чувствовать горький вкус, чем люди с аллелью G.

### Отличие между ассоциацией признака с аллелью и с генотипом

Например, цвет глаз

Eyes	Brown/ Other	Blue/ Green
rs7495174		
A	55	22
G	21	0
AA	22	11
AG	11	0
GG	5	0
AA / AG	16	0
GG / AG	5	0
AA / AG	33	11

Allelic P-value = 0.006 (значимо) -

аллель A статистически значимо ассоциируется с голубыми глазами

Genotype p-value (A - рецессивный) = 0.009

генотип AA статистически значимо ассоциируется с голубыми глазами

Genotype p-value (A - доминантный) = 0.205 (>0.05)

нет статистической значимости

**Вывод: Цвет глаз: rs7495174, генотип AA ассоциирован с голубыми глазами, при этом A - рецессивный аллель**

## Переносимость лактозы

Lactose	Intolerant	Tolerant
rs4988235		
A	5	26
G	23	44
AA	1	8
AG	3	10
GG	10	17

AA		1		8	
GG / AG		13		27	
GG		10		17	
AA / AG		4		18	
+-----+	+-----+		+-----+		+-----+

Allelic P-value = 0.064 (незначимо на уровне 0.05, но близко) -  
аллель G близко к значимости ассоциации по переносимости лактозы

Genotype p-value (G -доминантный) = 0.2

нет статистической значимости

Genotype p-value (G - рецессивный) =0.147 (>0.05)

нет статистической значимости

Шанс найти фактор риска в основной группе 0.588

Но отношение шансов = 2.647

(тоже есть онлайн-калькулятор - <http://medstatistic.ru/calculators/calcodds.html>)

## Increased risk

What is the likelihood of seeing a trait given a genotype compared to overall likelihood of seeing the trait in the population?

Изначальный шанс иметь непереносимость к лактозе = 14 (всего людей с непереносимостью) / 49 = 0.28

Для генотипа GG  $10/(10+17)=10/27=0.37$

Повышение риска для генотипа GG по сравнению с базовым риском  
 $0.37/0.28=1.32$

Смотря по времени - для закрепления можно разобрать другие аллели и генотипы -

[http://web.stanford.edu/class/gene210/files/exercises/2015/2015\\_gwas\\_tables.txt](http://web.stanford.edu/class/gene210/files/exercises/2015/2015_gwas_tables.txt) -

# Odds Ratio, Increased Risk

		P-value	OR	IR
Lactose Intolerance	rs4988235	.09	2.7	1.2
Eye Color	rs7495174	.0093	0	inf
Asparagus	rs4481887	.084	2.35	1.18
Bitter Taste	rs713598	.000498	0.22	0.519
Earwax	rs17822931	.004	4.6	2.6

## Extra

### Посмотреть пример результатов тестирования 23 and me

Например, об ощущении горького вкуса в продуктах -  
[https://permalinks.23andme.com/pdf/samplereport\\_traits.pdf](https://permalinks.23andme.com/pdf/samplereport_traits.pdf)

Genetics of bitter taste detection: The TAS2R38 gene contains instructions for a protein, or taste receptor, that can detect the bitter chemical called "PTC." PTC isn't usually found in the human diet, but it is similar to chemicals present in vegetables like broccoli and brussels sprouts. People with the G variant have a taste receptor that can detect these PTC-like chemicals. This means people with the G variant may taste bitterness in these foods and avoid them all together.

Посмотреть отчеты об ассоциации с заболеваниями

<https://www.23andme.com/dna-reports-list/>

[https://permalinks.23andme.com/pdf/samplereport\\_genetichealth.pdf](https://permalinks.23andme.com/pdf/samplereport_genetichealth.pdf)