

Медицинская геномика

Василий Евгеньевич Раменский
Анастасия Александровна Жарикова и Мария Ильинична Зайченко

ramensky@gmail.com, azharikova89@gmail.com

НМИЦ Терапии и профилактической медицины
Факультет биоинженерии и биоинформатики МГУ
Институт искусственного интеллекта МГУ

2024

Некоторые основы количественной генетики

- 1 Менделевское наследование: менделевцы vs биометрики
- 2 Дисперсия фенотипа и наследуемость
- 3 Наследуемость признаков у человека: примеры

Классы фенотипов

Дихотомические (бинарные) фенотипы

- Менделевские (моногенные) признаки

Комплексные (частые, мультифакторные) заболевания

- Диабет
- Шизофрения
- Коронарная (ишемическая) болезнь сердца (ИБС)

Количественные фенотипы

- Возраст
- Рост
- Индекс массы тела (ИМТ)
- Уровни липидов крови (ЛПНП, ЛПВП, ...)
- Артериальное давление крови (сАД, дАД)

Спор менделевцев и биометриков (1900-1918)

Предпосылки:

- 1865 Грегор Мендель: "Опыты над растительными гибридами" (Versuche über Pflanzen-Hybriden)
- 1865 Фрэнсис Гальтон "Наследственный талант и характер" (Hereditary Talent and Character)

Биометрики: большинство важных для эволюции характеристик (плодовитость, размер тела, сила, навык ловли добычи и сбора пищи, ...) являются непрерывными или количественными характеристиками и не поддаются Менделевскому анализу

Спор менделевцев и биометриков (1900-1918)

Предпосылки:

- 1865 Грегор Мендель: "Опыты над растительными гибридами" (Versuche über Pflanzen-Hybriden)
- 1865 Фрэнсис Гальтон "Наследственный талант и характер" (Hereditary Talent and Character)

Биометрики: большинство важных для эволюции характеристик (плодовитость, размер тела, сила, навык ловли добычи и сбора пищи, ...) являются непрерывными или количественными характеристиками и не поддаются Менделевскому анализу

Разрешение спора:

- 1918 R.A.Fisher, "The Correlation between Relatives on the Supposition of Mendelian Inheritance"
- 1965 D.S.Falconer, "The inheritance of liability to certain diseases, estimated from the incidence among relatives"



Спор менделевцев и биометриков (1900-1918)

XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. By R. A. Fisher, B.A. Communicated by Professor J. ARTHUR THOMSON. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)

CONTENTS.

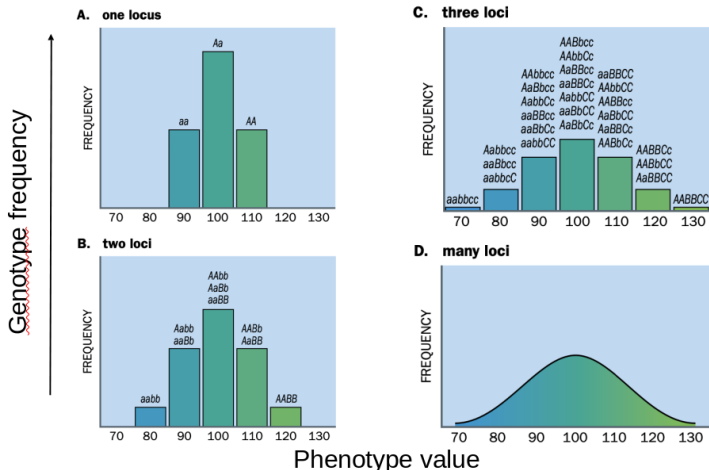
	PAGE		PAGE
1. The superposition of factors distributed independently	402	15. Homogamy and multiple allelomorphism	416
2. Phase frequency in each array	402	16. Coupling	418
		17. Theories of marital correlation; ancestral	---

Several attempts have already been made to interpret the well-established results of biometry in accordance with the Mendelian scheme of inheritance. It is here attempted to ascertain the biometrical properties of a population of a more general type than has hitherto been examined, inheritance in which follows this scheme. It is hoped that in this way it will be possible to make a more exact analysis of the causes of human variability. The great body of available statistics show us that the deviations of a human measurement from its mean follow very closely the Normal Law of Errors, and, therefore, that the variability may be uniformly measured by the standard deviation corresponding to the square root

3. Hence the members of this array mating at random will have offspring distributed in the three phases in the proportion

$$\begin{aligned}
 & \bar{P}^2 \left[1 + \frac{x}{\sigma^2}(a-m) \right] + \bar{P}\bar{Q} \left[2 + \frac{x}{\sigma^2}(a-m+d-m) \right] + \bar{Q}^2 \left[1 + \frac{x}{\sigma^2}(d-m) \right], \\
 & \bar{P}\bar{Q} \left[2 + \frac{x}{\sigma^2}(a-m+d-m) \right] + 2\bar{Q}^2 \left[1 + \frac{x}{\sigma^2}(d-m) \right] + \bar{P}\bar{R} \left[2 - \frac{x}{\sigma^2}(2m) \right] + \bar{Q}\bar{R} \left[2 + \frac{x}{\sigma^2}(d-m-a-m) \right],
 \end{aligned}$$

Полигенная природа количественных фенотипов



Гипотетическая непрерывная характеристика, среднее = 100 единиц, аддитивные (кодоминантные) эффекты: каждая копия A = +5 единиц, a = -5 единиц. Все частоты аллелей равны 0.5. // Strachan, Read – *Human Molecular Genetics*

Дисперсия фенотипа

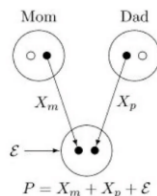
Значение фенотипа P : $P = X_m + X_p + \epsilon$,

где $P = P_{individual} - P_{pop.mean}$ – отклонение фенотипа индивидуума от популяционного среднего; $X_{m,p}$ (maternal, paternal) и ϵ – [нормально распределенные] случайные величины.

Обозначим дисперсию: $V(X_m) = V(X_p) = V_a/2$, $V(\epsilon) = V_e$
 J.Gillespie: «Количественная генетика – исключительно про дисперсию, ковариацию и корреляцию»

Дисперсия фенотипа: $V(p) = V(X_m) + V(X_p) + V(\epsilon) + 2C(X_m, X_p) + 2C(X_m, \epsilon) + 2C(X_p, \epsilon)$,
 где $V(\cdot)$ обозначает дисперсию, $C(\cdot, \cdot)$ – ковариацию, $C(X_m, X_p) = 0$, если родители не являются родственниками, $C(X, \epsilon) = 0$ – взаимодействие генотипа и окружающей среды

Заметьте: $C(X, X) = V(X) = V_a/2$



Gillespie – *Population genetics. A concise guide*

Пример взаимодействия генотипа и окружающей среды:

Один аллель прибавляет +1 см к фенотипу в теплой окружающей среде и -1см в холодной; второй аллель действует наоборот.

Наследуемость // Mousseau и Roff (1987)

$$V_p = V(X_m) + V(X_p) + V(\epsilon) + 2C(X_m, X_p) + 2C(X_m, \epsilon) + 2C(X_p, \epsilon) = V_a/2 + V_a/2 + V(\epsilon)$$

Дисперсия фенотипа является суммой **аддитивной** дисперсии и дисперсии окружающей среды:

$$V_p = V_a + V_e$$

Наследуемость [в узком смысле] признака h^2 :

$$h^2 = \frac{V_a}{V_p} = \frac{V_a}{V_a + V_e}$$

Аддитивная дисперсия = генетический вклад равняется сумме вклада от каждого аллеля (который не взаимодействует с другими)

Вид	Признак	Наследуемость
Пчела	Потребление кислорода	0.15
<i>Eurytemora herdmani</i>	Длина тела	0.12
Сверчок	Длина крыла	0.74
Хрущак мучной	Плодовитость	0.36
Саламандра	Количество позвонков	0.61
Дарвинов вьюрок	Вес	0.91
Дарвинов вьюрок	Длина клюва	0.85

Корреляция фенотипов

Фенотипы у случайной пары родственников:

$$P_X = X_m + X_p + \epsilon_x$$

$$P_Y = Y_m + Y_p + \epsilon_y$$

Поэтому:

$$\begin{aligned} C(P_X, P_Y) &= C(X_m, Y_m) + C(X_m, Y_p) + C(X_m, \epsilon_y) + C(X_p, Y_m) + C(X_p, Y_p) + \\ &\quad C(X_p, \epsilon_y) + C(\epsilon_x, Y_m) + C(\epsilon_x, Y_p) + C(\epsilon_x, \epsilon_y) \\ &= C(X_m, Y_m) + C(X_m, Y_p) + C(X_p, Y_m) + C(X_p, Y_p) \\ &= r_0 \times 0 + r_1 \times \frac{V_a}{2} + r_2 \times 2 \times \frac{V_a}{2} = rV_a \end{aligned}$$

где $C(X, X) = \frac{V_a}{2}$; вероятности того, что у двоих родственников 0, 1, 2 аллеля в состоянии IBD обозначены r_0, r_1, r_2 , а коэффициент родственности $r = \frac{r_1}{2} + r_2$

Родство	r_0	r_1	r_2	$r = \frac{r_1}{2} + r_2$
Родитель-потомок	0	1	0	1/2
Братья/сестры («siblings»)	1/4	1/2	1/4	1/2
Единокровные/единоутробные	1/2	1/2	0	1/4
Двоюродные	3/4	1/4	0	1/8

Корреляция фенотипов

$$C(P_X, P_Y) = rV_a$$

Помним, что $Corr(x, y) = \frac{C(x, y)}{\sqrt{V(x)V(y)}}$ и $V(X) = V_a/2$, поэтому:

$$Corr(P_X, P_Y) = \frac{C(P_X, P_Y)}{V_p} = \frac{rV_a}{V_p} = rh^2$$

$$Corr(P_X, P_Y) = rh^2$$

Корреляция между [фенотипическими значениями] пары родственников равняется произведению коэффициента родственности и наследуемости.

Зная отклонение фенотипа индивидуума X от популяционного среднего $P_X = x$, ожидаемое значение фенотипа родственника Y :

$$E(P_Y | P_X = x) = Corr(P_X, P_Y)x = rh^2x$$

Упражнение

Какова корреляция между каждой парой родственников в таблице выше, если $V_a = 2$, $V_p = 3$?

Наследуемость в широком и в узком смысле

Ранее: упрощенная модель только с аддитивными эффектами:

$$P = X_m + X_p + \epsilon$$

$$V_p = V_a + V_e$$

Более реалистично:

$$P = X_m + X_p + X_{mp} + \epsilon$$

где X_{mp} отражает доминантные отношения между материнским и отцовским аллелями. Предполагая, что аддитивный и доминантный вклады не коррелируют:

$$V_p = V_a + V_d + V_e$$

где V_d так называемая **дисперсия доминантности**. Если мы учитываем несколько взаимодействующих локусов, тогда

$$V_p = V_a + V_d + V_i + V_e$$

где V_i отражает **дисперсию эпистаза** из-за взаимодействия между локусами.

Наследуемость в широком и в узком смысле

Наследуемость [в узком смысле] h^2 все еще

$$h^2 = \frac{V_a}{V_p}$$

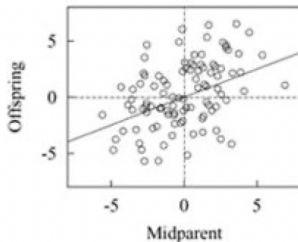
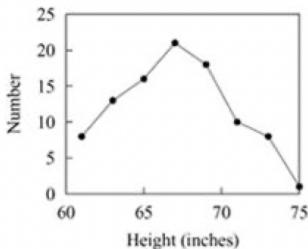
но наследуемость в широком смысле H^2 является отношением всех генетических дисперсий к фенотипической дисперсии

$$H^2 = \frac{V_a + V_d + V_i}{V_p}$$

Упражнение

В популяции жуков общая дисперсия веса $V_p = 130$; дисперсия факторов окружающей среды $V_e = 35$ и доминантная генетическая дисперсия $V_d = 45$. Предполагая отсутствие эпистаза, рассчитайте наследуемость в узком смысле.

Пример признака: рост человека // Strachan, Read – *Human Molecular Genetics*



Гистограмма и отклонение от популяционного среднего, в дюймах

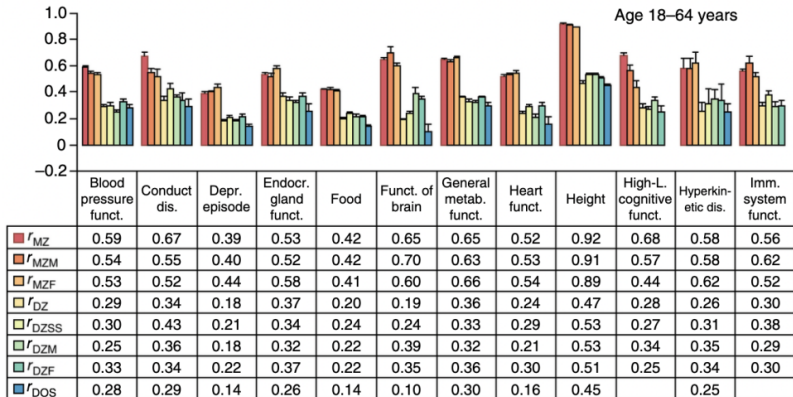
- Коэффициент корреляции $r^2 = 0.476 > 0 \Rightarrow$ родственники напоминают друг друга \Rightarrow у признака есть генетическая компонента
- Коэффициент $r^2 < 1 \Rightarrow$ «регрессия к среднему»
- Разброс из-за Менделевской сегрегации и эффектов окружающей среды

Meta-analysis of the heritability of human traits based on fifty years of twin studies

Tinca J C Polderman^{1,10}, Beben Benyamin^{2,10}, Christiaan A de Leeuw^{1,3}, Patrick F Sullivan⁴⁻⁶, Arjen van Bochoven⁷, Peter M Visscher^{2,8,11} & Danielle Posthuma^{1,9,11}

702

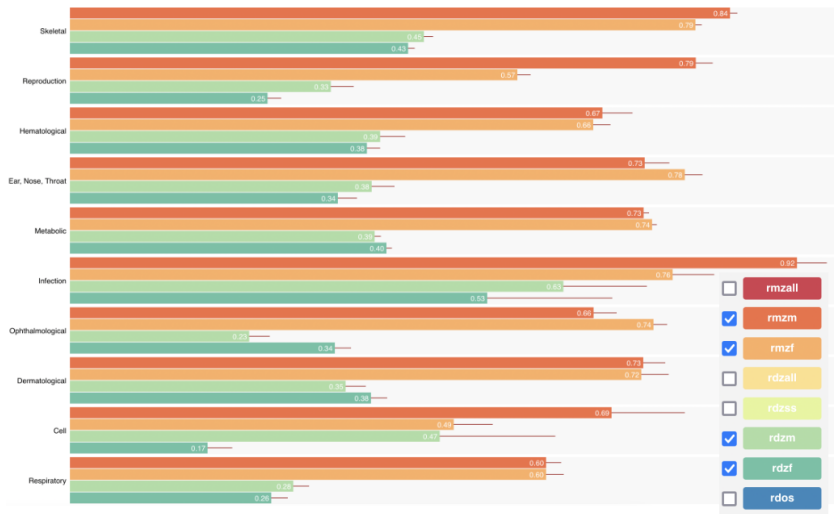
VOLUME 47 | NUMBER 7 | JULY 2015 NATURE GENETICS



MaTCH

Meta-Analysis of Twin Correlations and Heritability

This website provides a resource for the heritability of all human traits that have been investigated with the classical twin design.



https://match.ctglab.nl/#/multiple/twin_cor

Heritability of 596 lipid species and genetic correlation with cardiovascular traits in the Busselton Family Heart Study[§]

Gemma Cadby,^{1,*†} Phillip E. Melton,^{†,§***} Nina S. McCarthy,[†] Corey Giles,^{††} Natalie A. Mellett,^{††} Kevin Huynh,^{††} Joseph Hung,^{§§****} John Beilby,^{††,§§§} Marie-Pierre Dubé,^{*****} Gerald F. Watts,^{§§,†††} John Blangero,^{§§§§} Peter J. Meikle,^{2,††} and Eric K. Moses^{2,†,§}

Journal of Lipid Research Volume 61, 2020 **537**

TABLE 1. Characteristics of study population and heritability of CVD traits

	Mean (SD), n = 4,492	h ² (SE)
Age, years	50.83 (17.37)	—
BMI, ^a kg/m ²	26.04 (4.23)	0.46 (0.04)
WHR ^a	0.85 (0.07)	0.25 (0.03)
HDL-C, ^b mmol/l	1.39 (0.39)	0.59 (0.03)
LDL-C, ^b mmol/l	3.60 (1.00)	0.52 (0.04)
Triglycerides, ^b mmol/l	1.32 (0.93)	0.37 (0.03)
Total cholesterol, ^b mmol/l	5.59 (1.11)	0.57 (0.03)
SBP, ^c mmHg	124.6 (19.33)	0.32 (0.04)
DBP, ^c mmHg	75.09 (10.72)	0.26 (0.04)