

Fine mapping and functional analysis

Семинар наставника 07.10.2024

Мичил Трофимов



Повтор: Genome-wide Association Study (GWAS)

Основная задача Полно-геномного поиска ассоциаций (GWAS) -- найти ассоциации фенотипа с однонуклеотидными полиморфизмами (SNP) во всём геноме

- SNP = замена одного нуклеотида в ДНК (пример: A → T)

Повтор: Genome-wide Association Study (GWAS)

Основная задача Полно-геномного поиска ассоциаций (GWAS) -- найти ассоциации фенотипа с однонуклеотидными полиморфизмами (SNP) во всём геноме

- SNP = замена одного нуклеотида в ДНК (пример: A \rightarrow T)
- Большинство пар оснований будут инвариантными в разных популяциях
- Некоторые статистически пере- или недообогатены в определенной популяции по сравнению с «нормальной» популяцией

Повтор: Genome-wide Association Study (GWAS)

Основная задача Полно-геномного поиска ассоциаций (GWAS) -- найти ассоциации фенотипа с однонуклеотидными полиморфизмами (SNP) во всём геноме

- SNP = замена одного нуклеотида в ДНК (пример: A \rightarrow T)
- Большинство пар оснований будут инвариантными в разных популяциях
- Некоторые статистически пере- или недообогатены в определенной популяции по сравнению с «нормальной» популяцией
- Определенные популяции могут включать людей с определенным заболеванием, людей, которые негативно реагируют на определенное терапевтическое средство и другие фенотипические классификации

Повтор: Genome-wide Association Study (GWAS)

Основная задача Полно-геномного поиска ассоциаций (GWAS) -- найти ассоциации фенотипа с однонуклеотидными полиморфизмами (SNP) во всём геноме

- SNP = замена одного нуклеотида в ДНК (пример: A → T)
- Большинство пар оснований будут инвариантными в разных популяциях
- Некоторые статистически пере- или недообогатены в определенной популяции по сравнению с «нормальной» популяцией
- Определенные популяции могут включать людей с определенным заболеванием, людей, которые негативно реагируют на определенное терапевтическое средство и другие фенотипические классификации

«Библиотека» SNP, которые необходимо проверить, часто составляет порядка 600-800 тыс.

Повтор: Genome-wide Association Study (GWAS)

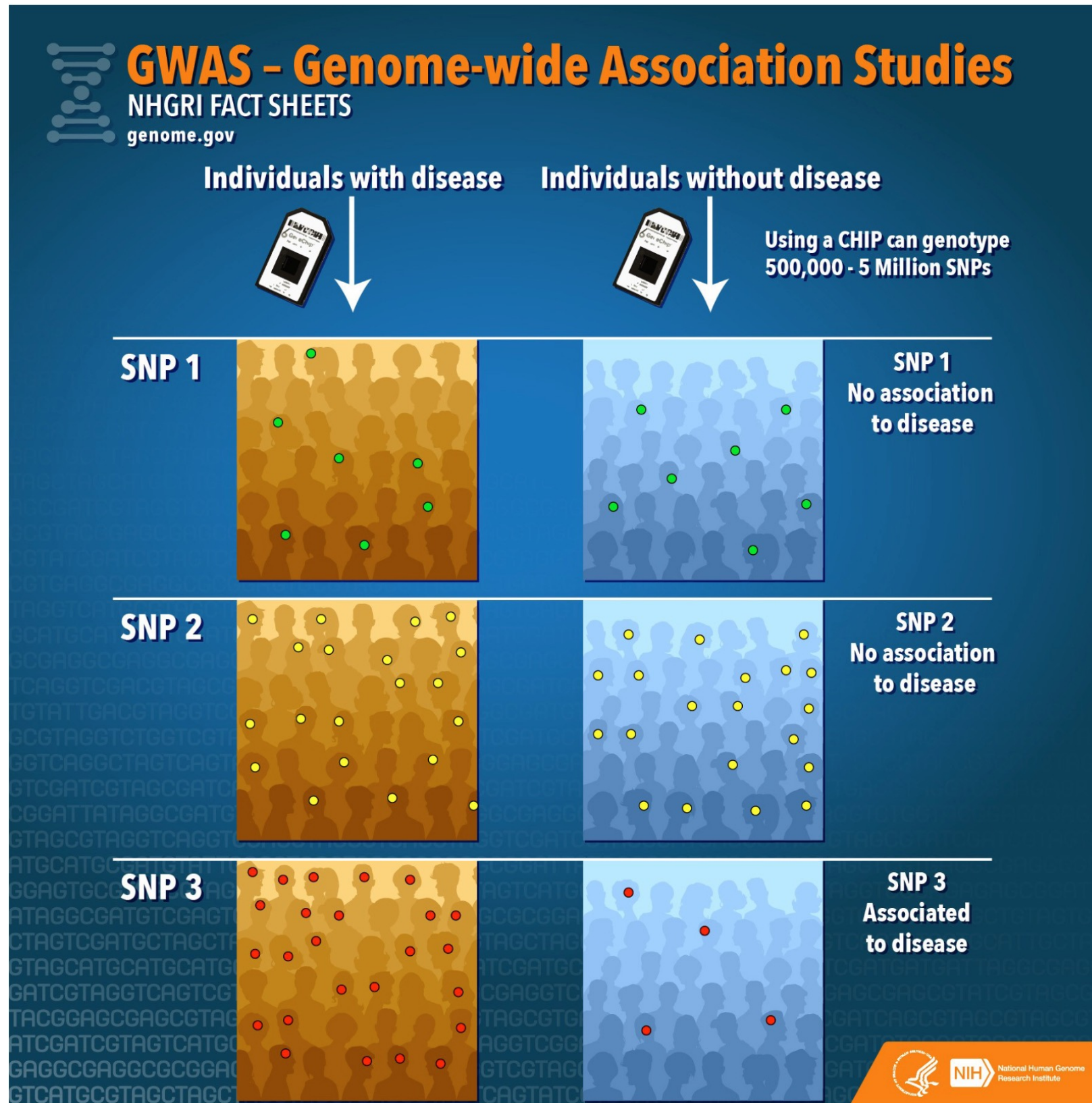
Основная задача Полно-геномного поиска ассоциаций (GWAS) -- найти ассоциации фенотипа с однонуклеотидными полиморфизмами (SNP) во всём геноме

- SNP = замена одного нуклеотида в ДНК (пример: A → T)
- Большинство пар оснований будут инвариантными в разных популяциях
- Некоторые статистически пере- или недообогатены в определенной популяции по сравнению с «нормальной» популяцией
- Определенные популяции могут включать людей с определенным заболеванием, людей, которые негативно реагируют на определенное терапевтическое средство и другие фенотипические классификации

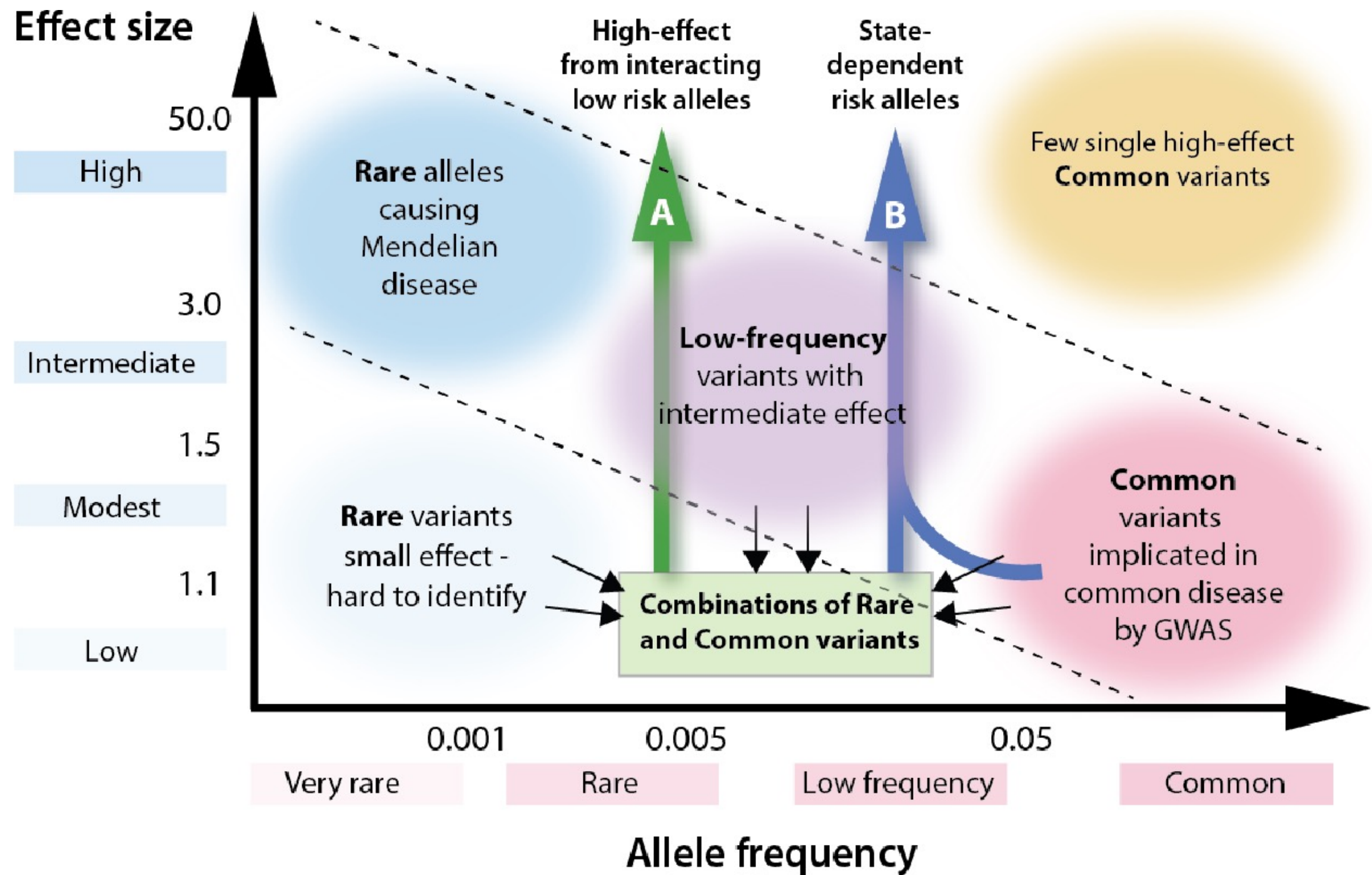
«Библиотека» SNP, которые необходимо проверить, часто составляет порядка 600-800 тыс.

- GWAS имеет большие размеры когорт (от тысяч до сотен тысяч) и хорошее покрытие участков
- Поэтому GWAS хорошо подходит для идентификации многих SNP, связанных с «полигенными» расстройствами, где каждый локус небольшое, но значимое влияние на фенотипическое изменение

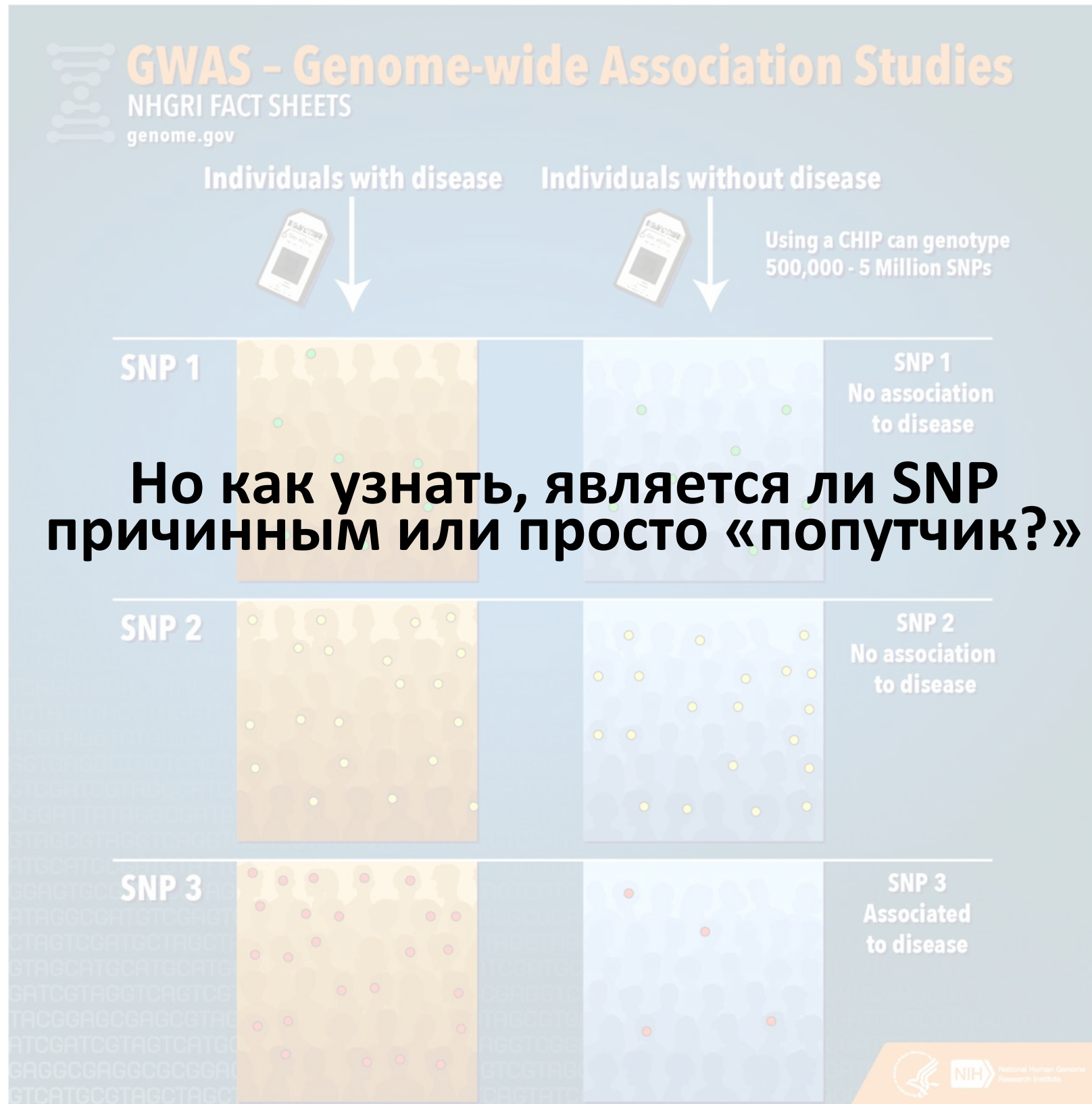
Повтор: Genome-wide Association Study (GWAS)



Повтор: моногенные и полигенные заболевания

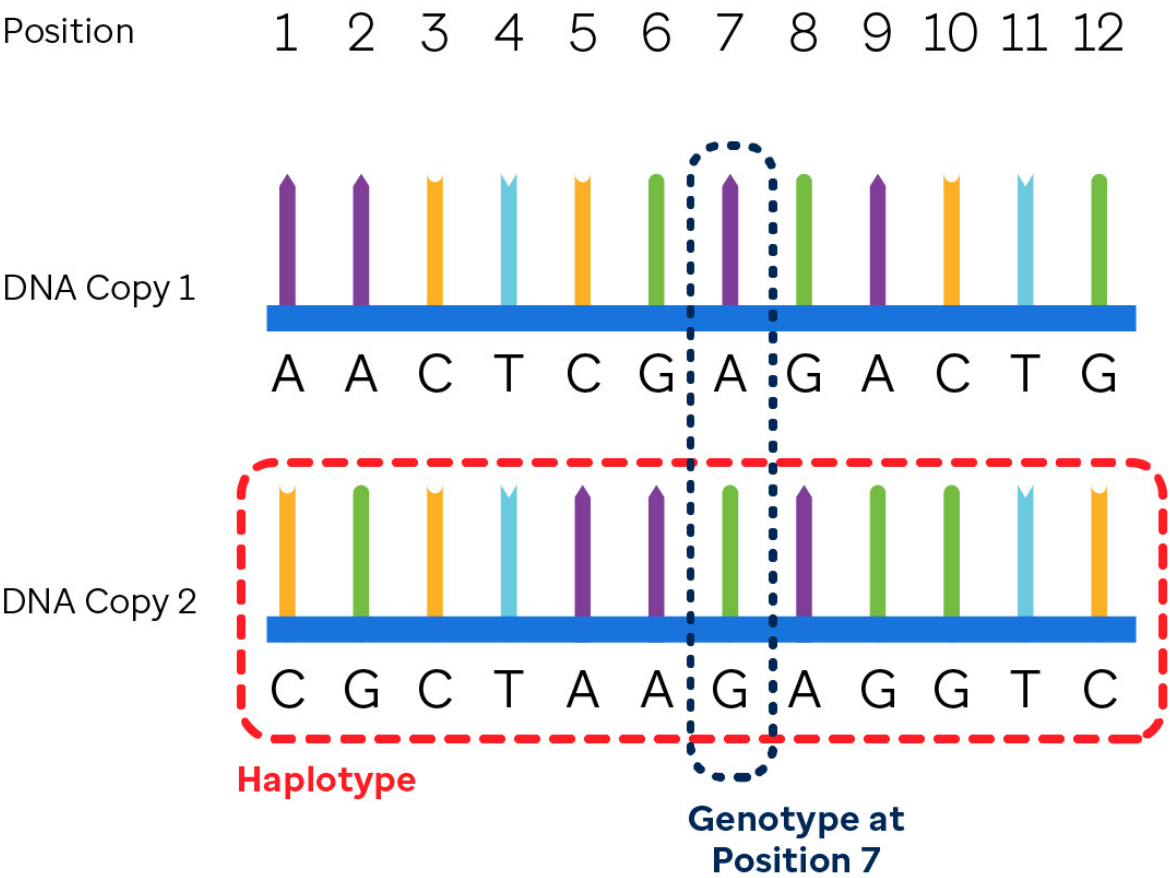


Повтор: Genome-wide Association Study (GWAS)

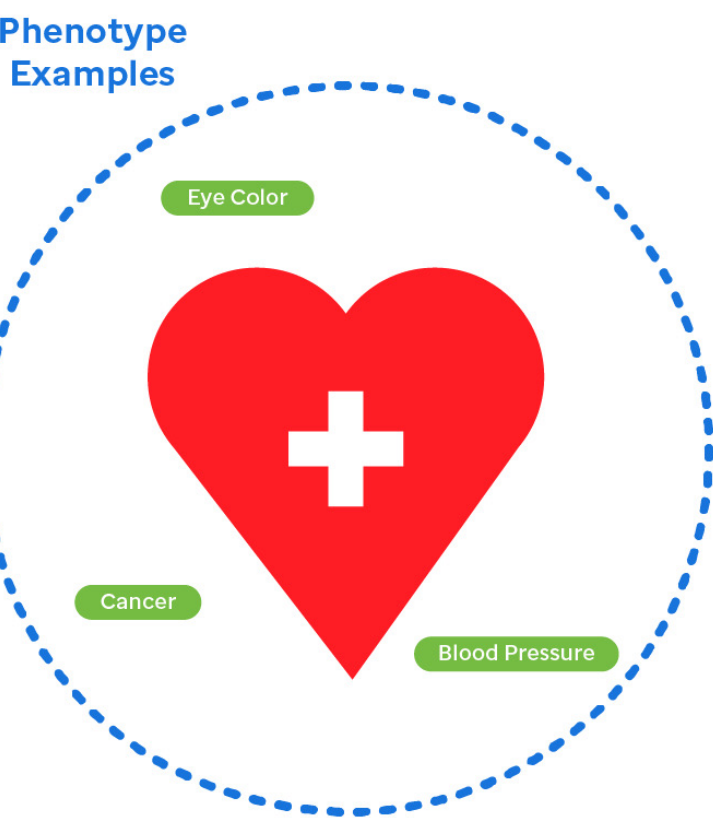


Повтор: гаплотипы и linkage disequilibrium (LD)

Genomic Sequence



Physical Effects



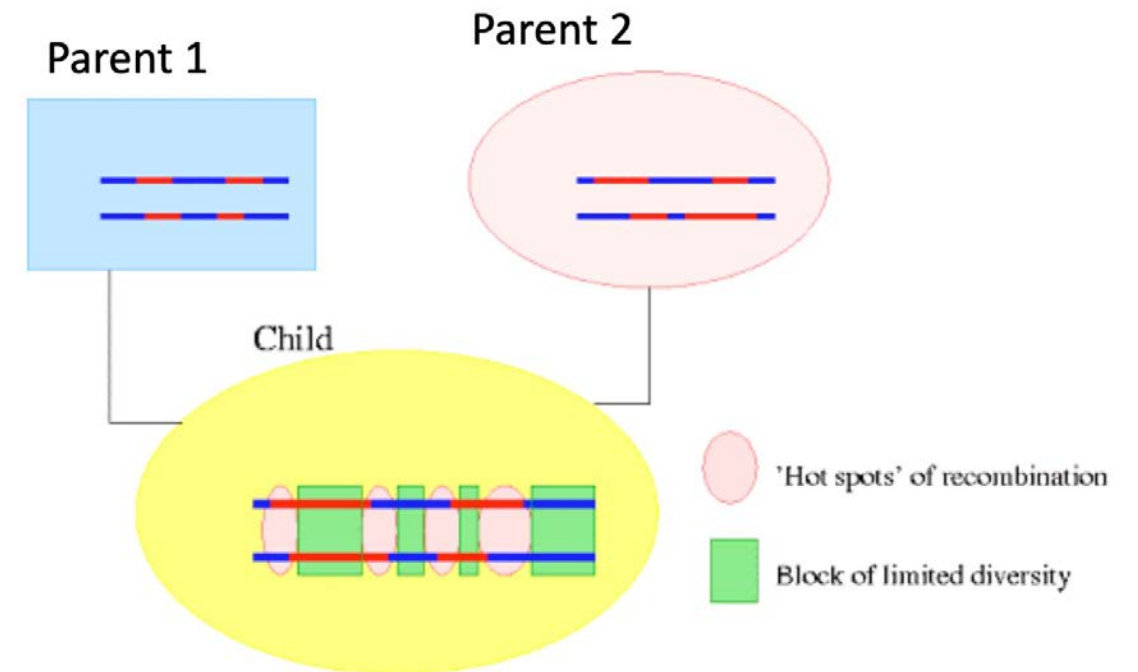
Повтор: гаплотипы и linkage disequilibrium (LD)

Аллельные группы, называемые «гаплотипами», являются важными искажающими факторами в анализе GWAS.

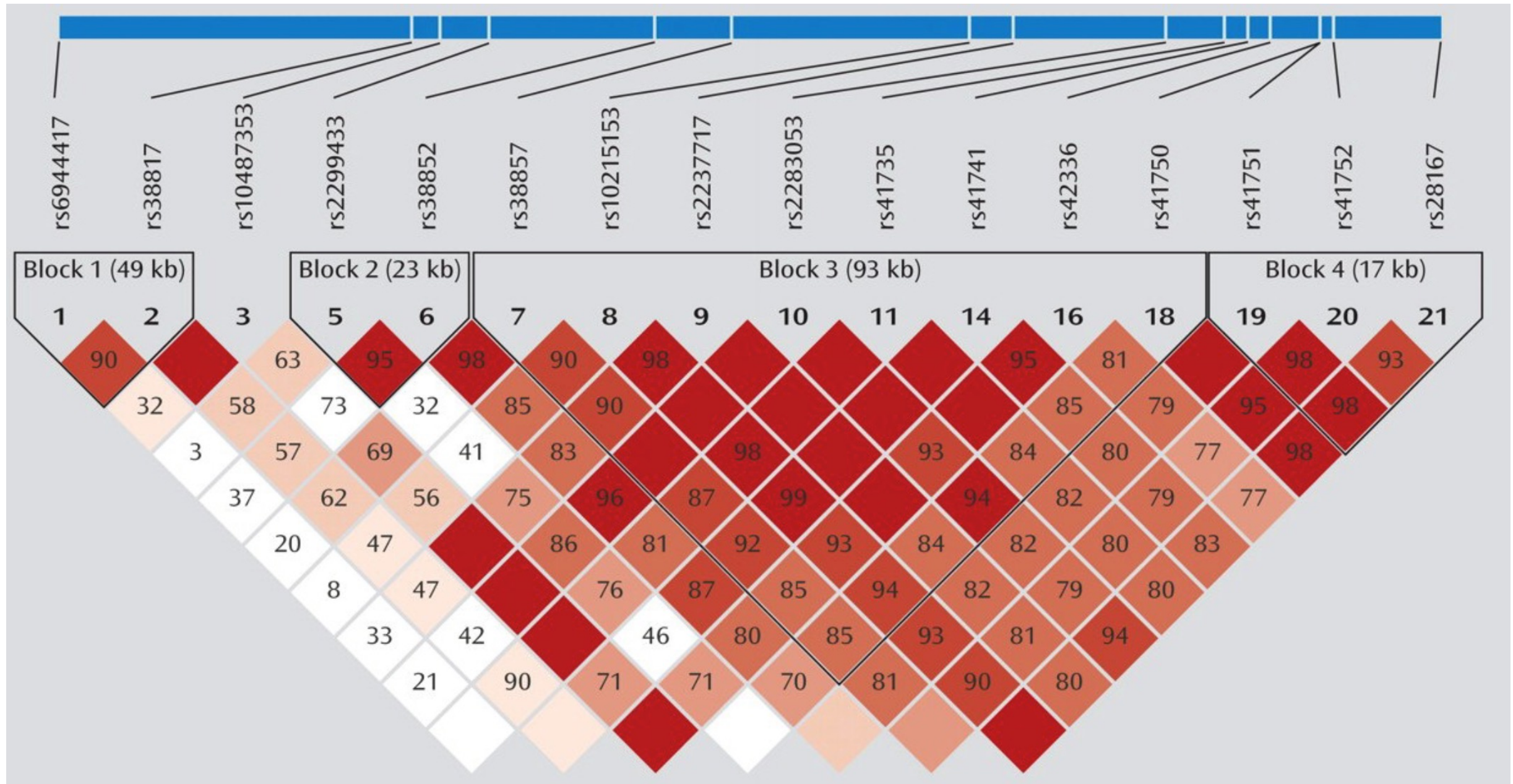
Гаплотипы возникают из-за рекомбинации

Низкое разнообразие гаплотипов может вызвать высокий LD = целый набор SNP, которые, как правило, возникают в определенных шаблонах

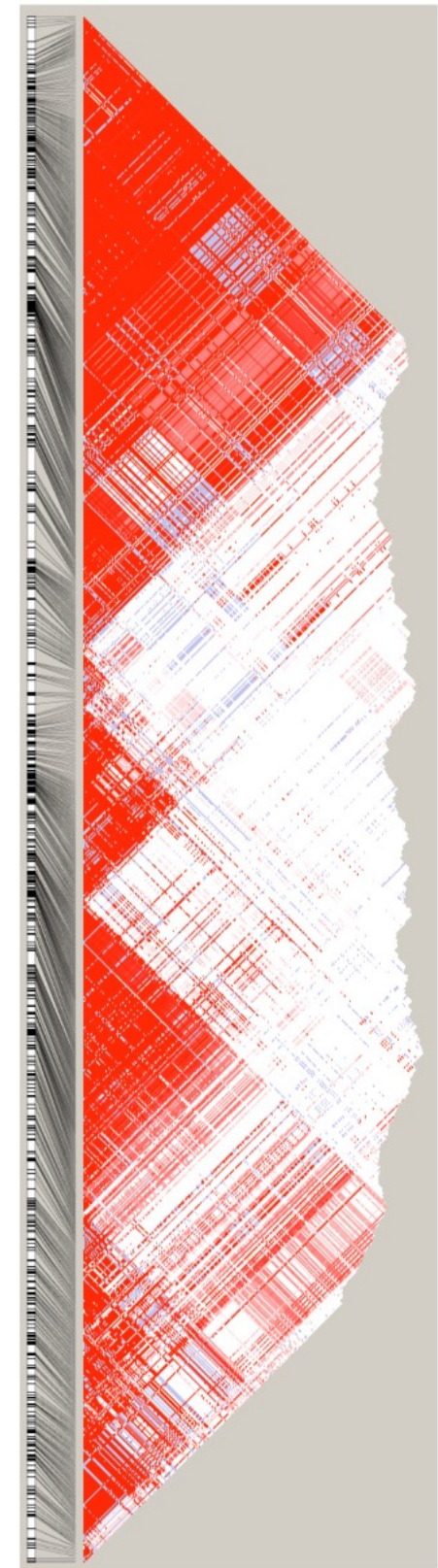
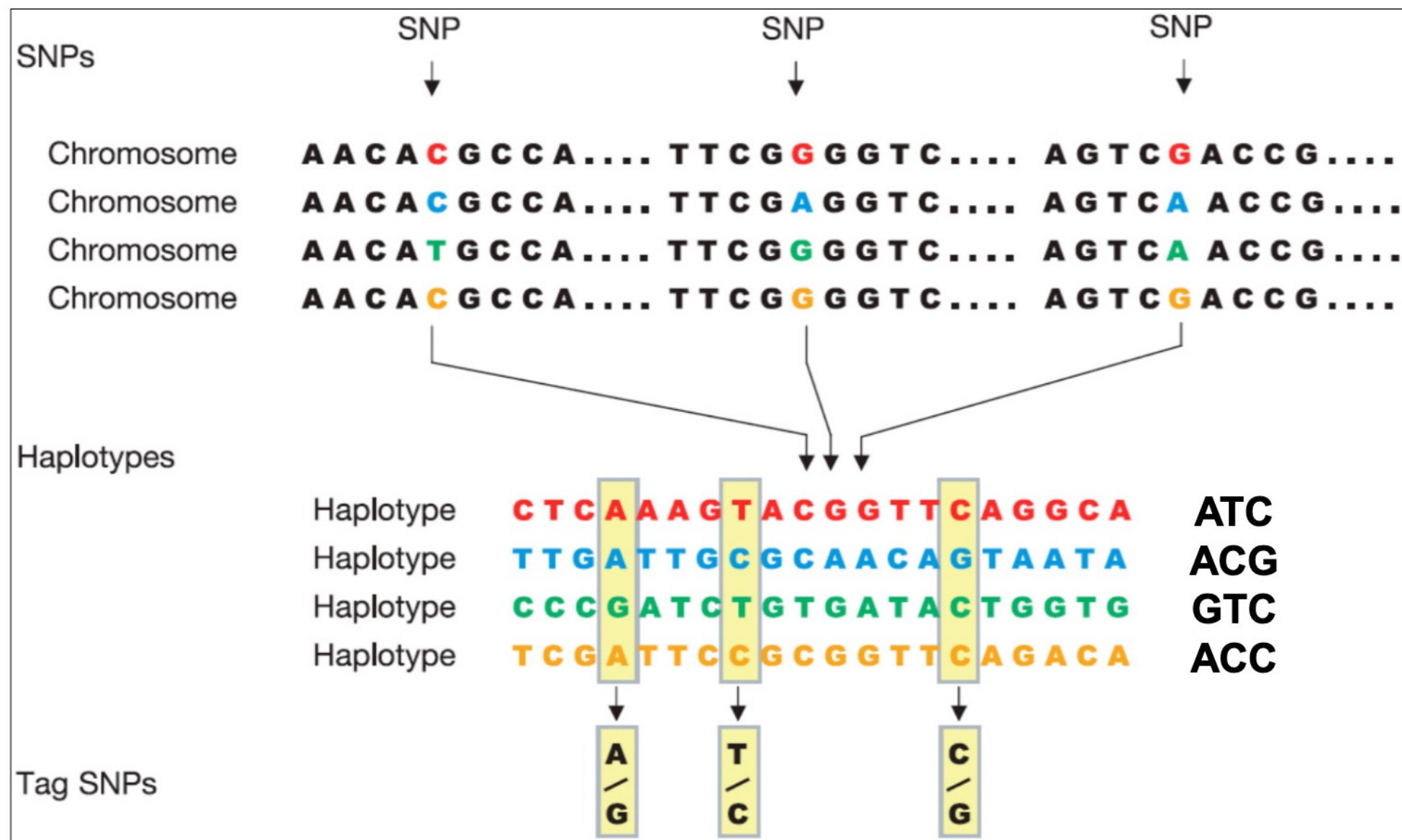
LD затрудняет разделение биологически значимых SNP-сайтов из тех, которые просто случайно включены в гаплотип!



Повтор: гаплотипы и linkage disequilibrium (LD)

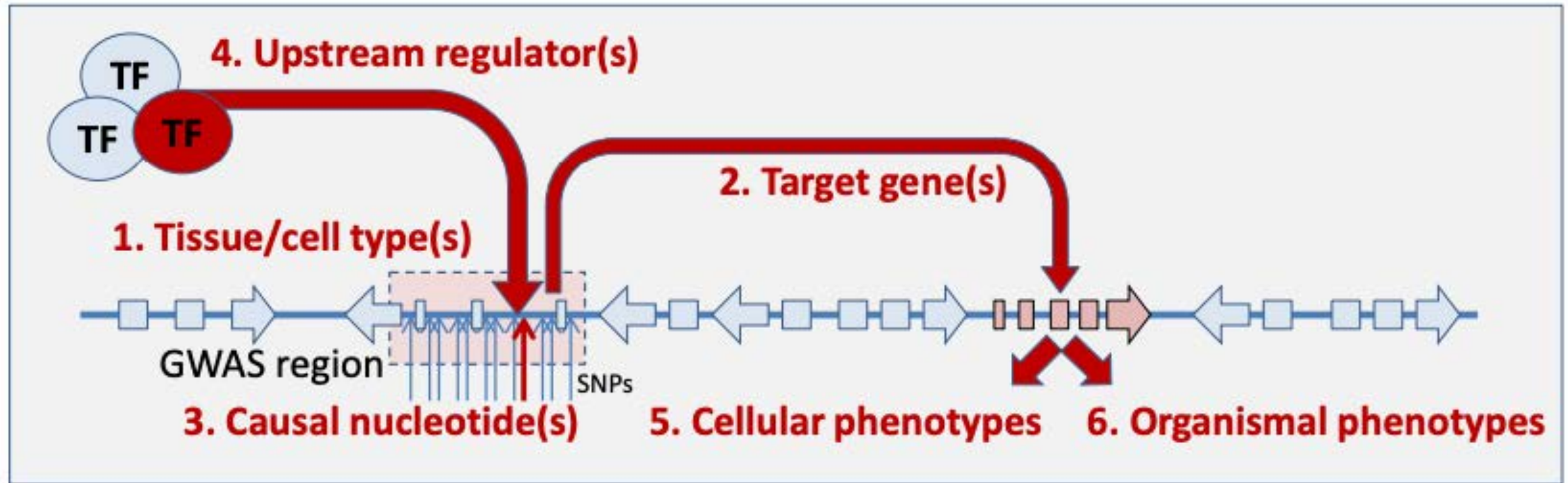


Повтор: распространенные SNP находятся в гаплотипах



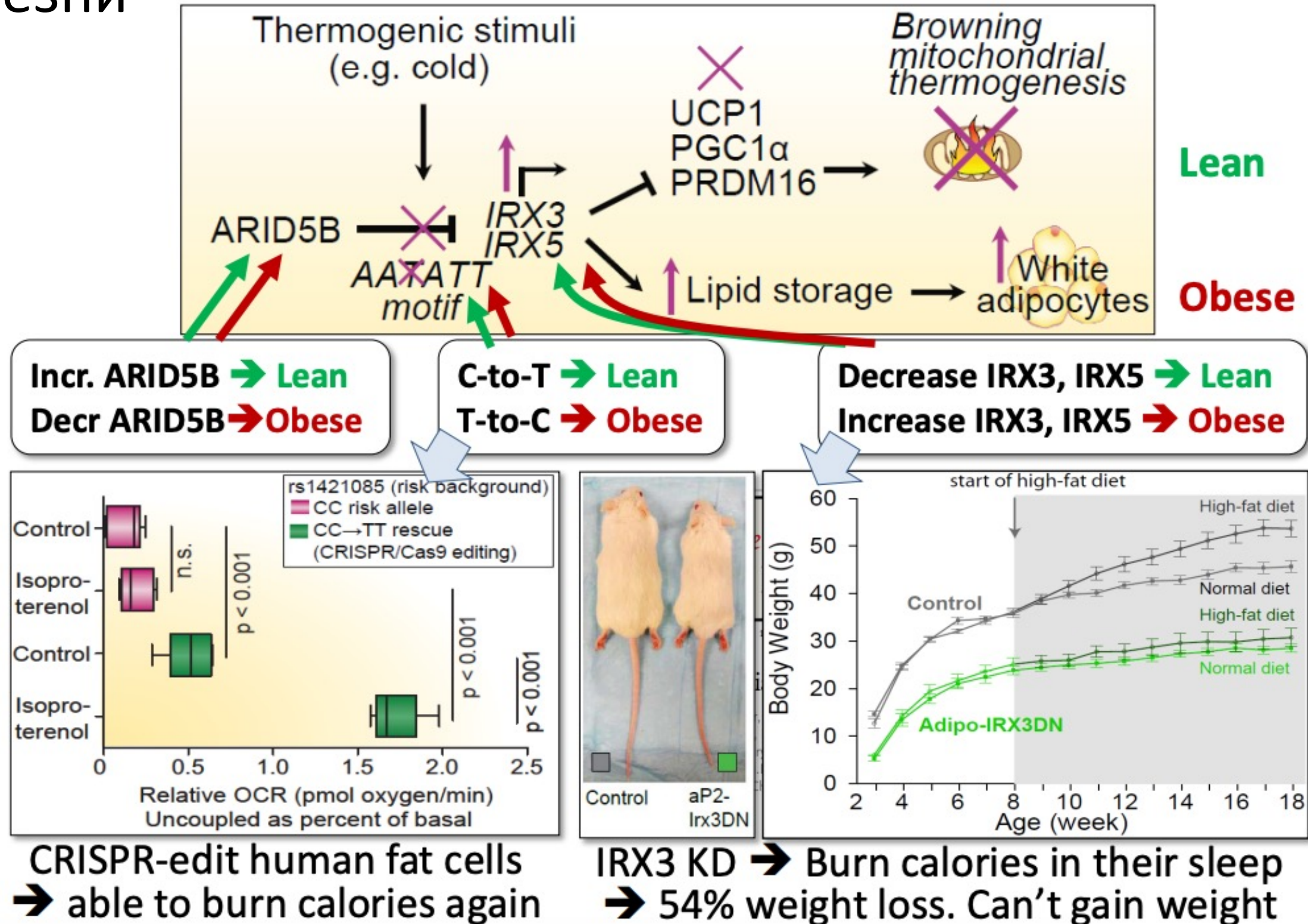
- распространенные SNP приходятся один раз на каждые 1000 нуклеотидов
- Они наследуются вместе, поэтому нужно только профилировать подмножество
- Маркеры, выбранные для профилирования гаплотипа, являются «теговыми» SNP

Ассоциация фенотипа с SNP в некодирующей области



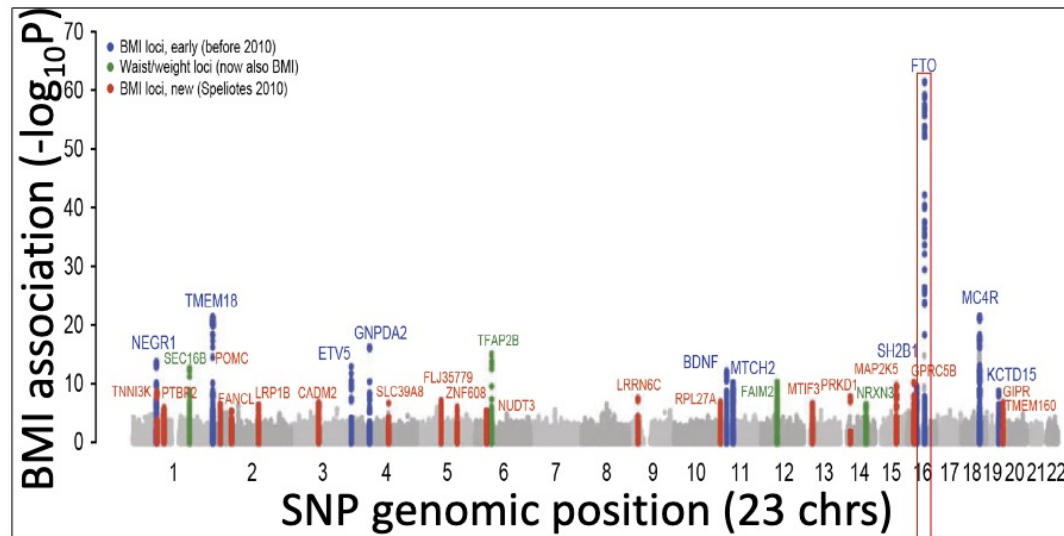
1. Установите тип соответствующей ткани/клетки
2. Определите целевой ген(ы), находящийся downstream от SNP
3. Установление причинного SNP
4. Установление TF
5. Установление клеточных фенотипических
6. Установление фенотипических последствий для организма

Установить биологическую сеть → обратить фенотип болезни

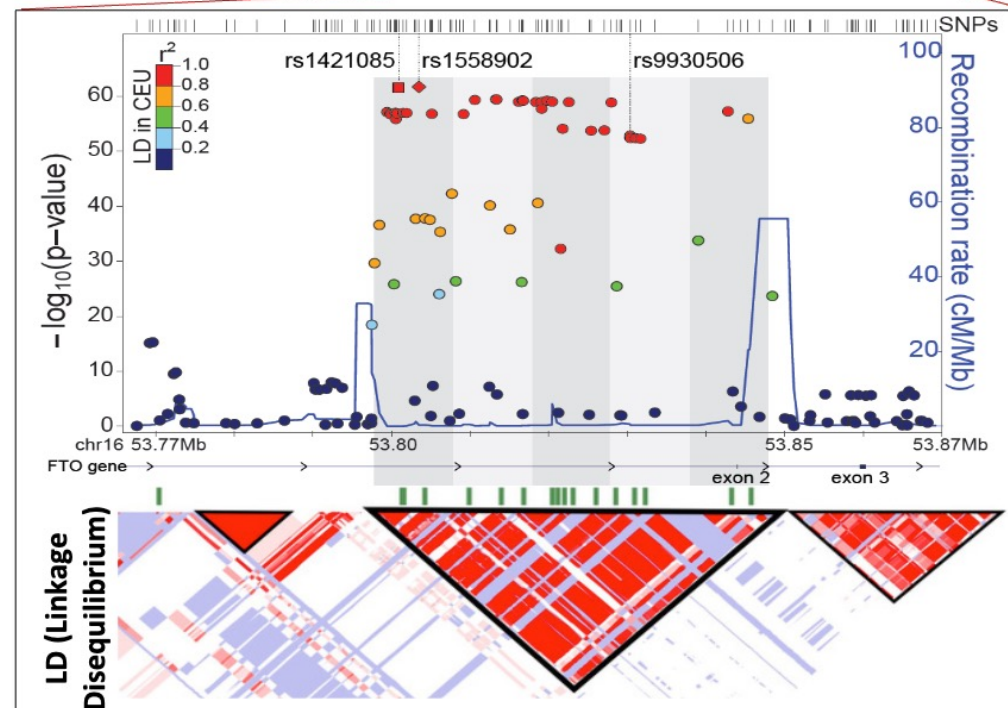


Genomic medicine today: challenge and promises

GWAS Manhattan Plot: simple χ^2 statistical test



Speliotes NG 2010



Dina NG 2007, Frayling Science 2007, Claussnitzer NEJM 2015

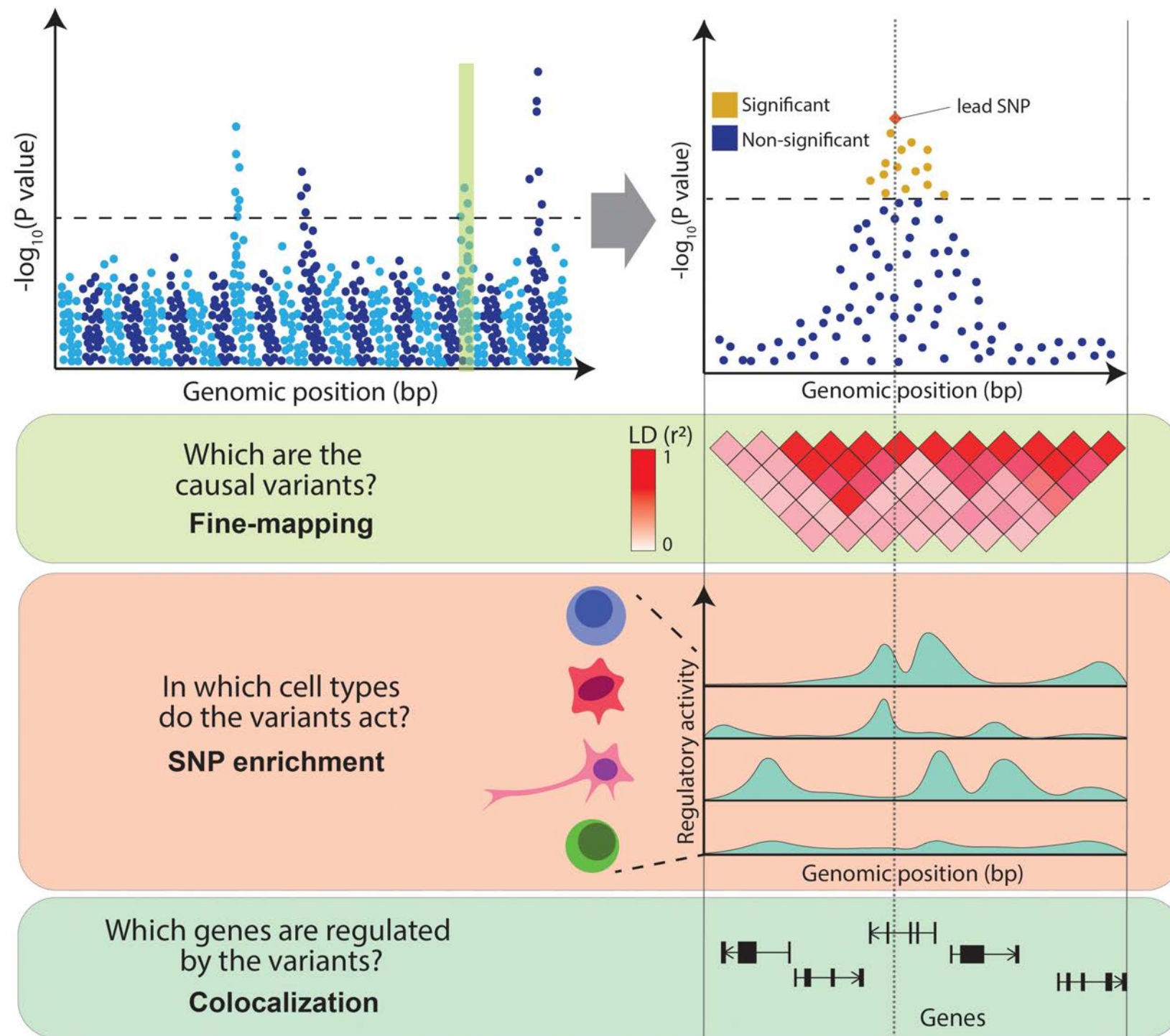
The promise of genetics

- Unbiased, Causal, Uncorrected
- New disease mechanisms
- New target genes
- New therapeutics
- Personalized medicine

The challenge of mechanism

- **90+% disease hits non-coding**
- Target gene not known
- Causal variant not known
- Cell type of action not known
- Relevant pathways not known
- Mechanism not known

Fine mapping и functional analysis причинного SNP



SusieR

SusieR: Sum of Single Effects Linear Regression

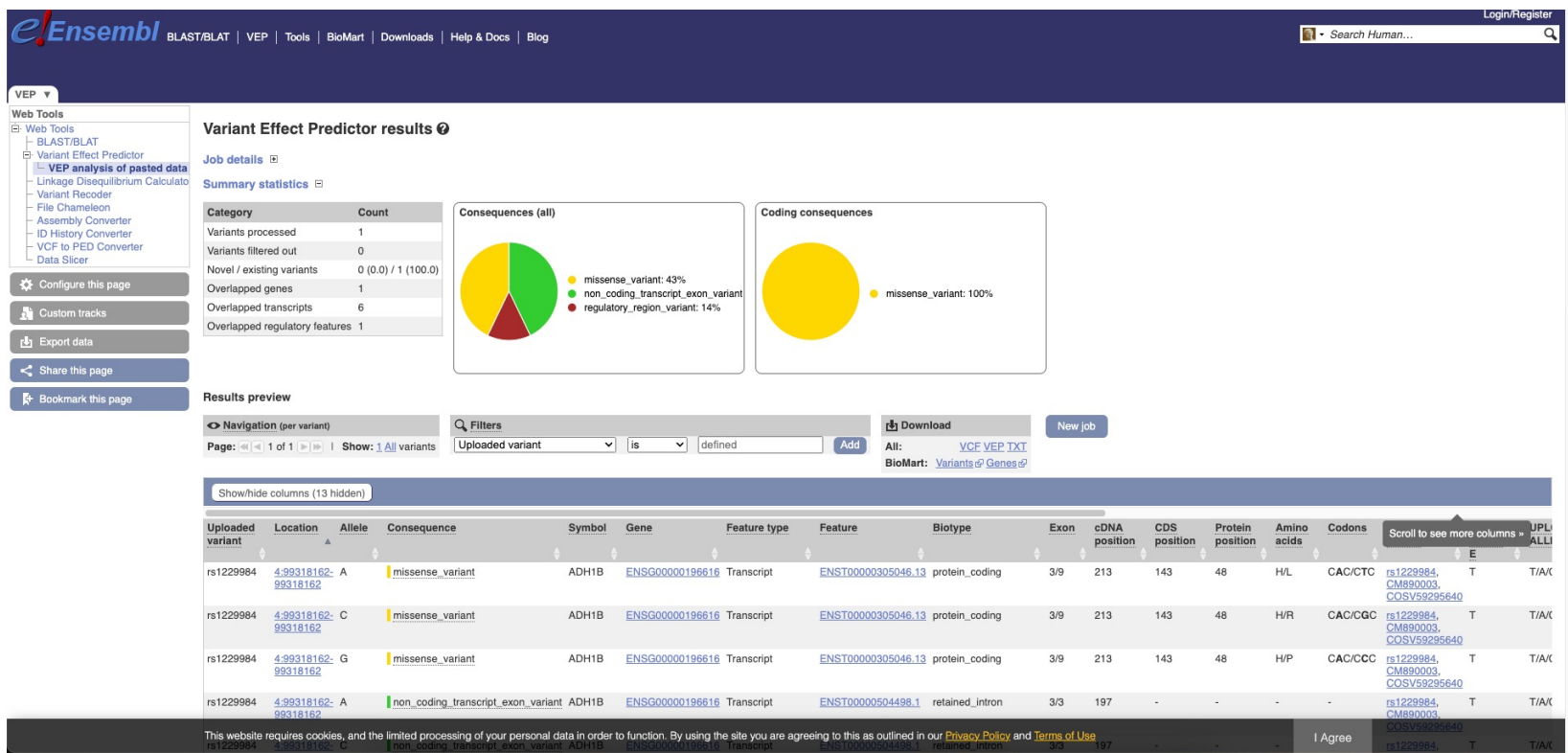
- Алгоритм «итеративного байесовского пошагового отбора»
- Производит «достоверные наборы» (CS) переменных, которые, вероятно, имеют ненулевые эффекты

https://stephenslab.github.io/susieR/articles/finemapping_summary_statistics.html

Functional analysis

После обнаружения причинного варианта, необходимо осветить следующие пункты:

- патогенность (Ensemble variant effect predictor: VEP)

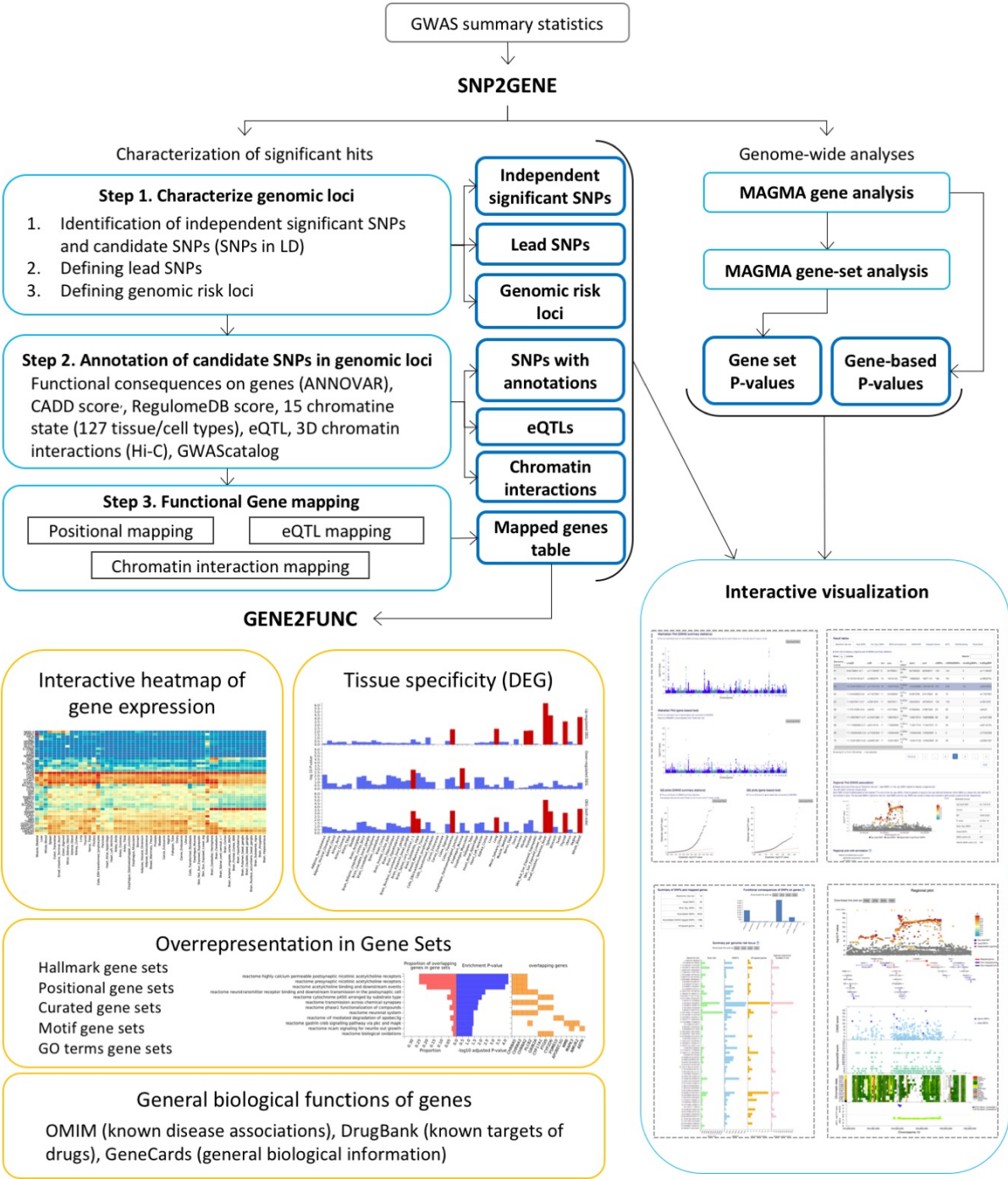


rs671

rs121913529

Functional analysis: FUMA

Функции гена в котором находится вариант, биологические процессы и сигнальные пути в которых он участвует



Functional analysis

Функции гена в котором находится вариант, биологические процессы и сигнальные пути в которых он участвует

FUMA

The screenshot shows the FUMA GWAS website. The navigation bar at the top includes links for Home, Tutorial, SNP2GENE, GENE2FUNC, Links, Login, and Register. The 'Login' and 'Register' links are highlighted with red boxes. Arrows point from these boxes to labels: '2) Login' points to the 'Login' box, and '1) Register' points to the 'Register' box. The main content area features the title 'FUMA GWAS' and the subtitle 'Functional Mapping and Annotation of genome-wide association results'. Below this, a paragraph describes the platform's purpose and the functions of SNP2GENE and GENE2FUNC. At the bottom, there are two preview panels: 'SNP2GENE' on the left and 'GENE2FUNC' on the right, each showing sample output plots and tables.

FUMAGWAS

Home Tutorial **SNP2GENE** GENE2FUNC Links **Login** **Register**

FUMA GWAS

Functional Mapping and Annotation of genome-wide association results

FUMA is a platform that can be used to annotate, prioritize and visualize and interpret GWAS results.
The **SNP2GENE** function takes GWAS summary statistics or a list of rsid's as input, and provides extensive functional annotation for all SNPs in genomic areas identified by lead SNPs.
The **GENE2FUNC** function takes a list of geneids (as identified by SNP2GENE or as provided manually) and annotates genes in biological context
Please log in to use FUMA. If you have't registered yet, you can do from [here](#).
When using FUMA, please acknowledge Watanabe et al. xxx

SNP2GENE

GENE2FUNC

login: trofimov.michil@gmail.com

pass: PB.h@pPNf37iJ4v