



**Faculdade**  
**UNIMED**

# ESQUIZOFRENIA E PSICOSE - A GENÉTICA DOS TRANSTORNOS PSIQUIÁTRICOS

**AULA TRANSCRITA**

**SINTIA BELANGEIRO**

**PÓS-GRADUAÇÃO ONLINE**  
**SAÚDE MENTAL E ATENÇÃO PSICOSSOCIAL**

**PÓS-GRADUAÇÃO ONLINE**  
**SAÚDE MENTAL E ATENÇÃO PSICOSSOCIAL**

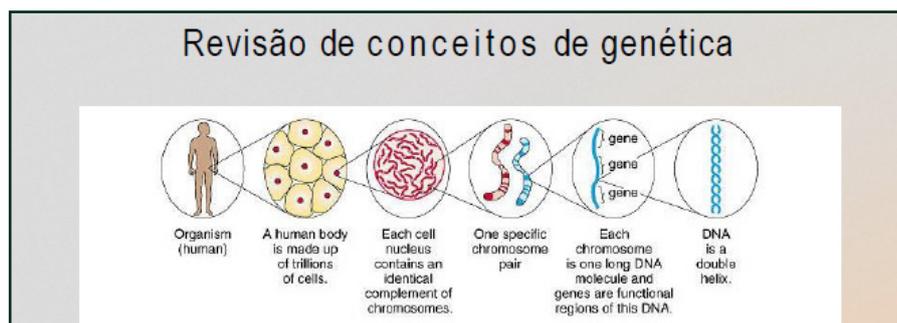
**Faculdade**  
**UNIMED**

## AULA 01

### APRESENTAÇÃO

Olá, eu sou a SÍntia Belangero, geneticista, professora da Universidade Federal de São Paulo, no Departamento de Morfologia e Genética. Hoje eu vou falar para vocês o respeito da Genética Psiquiátrica.

### REVISANDO CONCEITOS DE GENÉTICA



Antes de a gente falar um pouco sobre Psiquiatria, a gente vai conversar um pouquinho com uma revisão de conceitos em Genética. Então, como vocês já devem saber, nosso organismo é formado por trilhões de células e cada uma dessas células contém, dentro de seu núcleo, alguns cromossomos - no caso da espécie humana são 46 cromossomos humanos - e **dentro desses cromossomos humanos nós temos uma série de genes organizados cromossomo a cromossomo.**

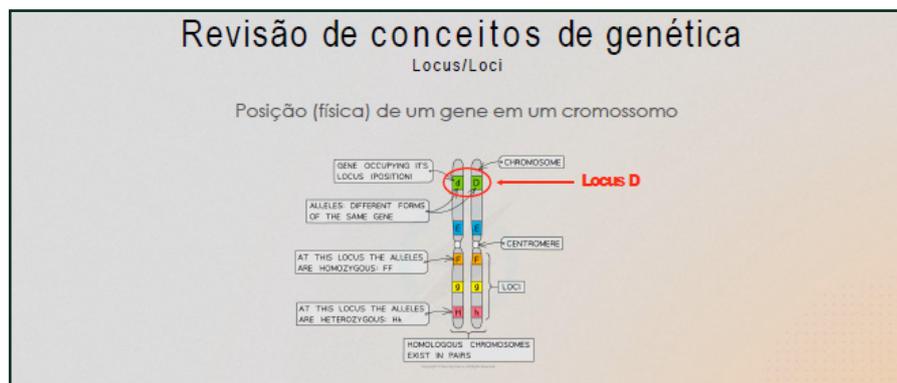


Esses genes são compostos de DNA. Esses cromossomos atingem o seu mais alto grau de compactação durante a metáfase, que é uma das fases da mitose, e esses cromossomos conseguem atingir esse mais alto grau de compactação com ajuda de algumas proteínas, chamadas **Histonas**. Elas são **proteínas básicas que estão ali presentes e disponíveis dentro do núcleo** e que são utilizadas como certo arcabouço para o DNA se compactar. É como se a gente tivesse alguns fios de lã e esses fios de lã estivessem enrolados num arcabouço, como uma analogia, poderia ser um carretel de linha. Essa estrutura de proteínas está dando **suporte para esse DNA se compactar**. E isso acontece durante a metáfase para que o material genético possa se reorganizar de forma a **se preparar para a próxima divisão celular**.

Nesse momento a gente tem estruturas **dentro do núcleo** que são chamadas de **cromossomos** e esses cromossomos são **formados por moléculas de DNA**, com ajuda de proteínas histonas.



Como a gente falou, esse cromossomo composto de uma molécula de DNA, cada um deles composto por uma molécula, essa molécula está subdividida em segmentos que são **sequências gênicas e não-gênicas**. Os **genes são sequências de DNA que codificam algum produto funcional**; esse produto funcional pode ser tanto um RNA como pode ser um RNA mensageiro, que depois vai ser codificado em forma de uma proteína. Dessa forma, a gente tem o conceito atual de gene, que seria **a sequência de um DNA que codifica um produto funcional**, e esse funcional pode ser um RNA ou então um RNA que depois será codificado em proteínas. Mas lembrando que **nem tudo dentro dos cromossomos são genes**, algumas partes do DNA, algumas sequências de DNA, são sequências que não codificam genes.



E ainda pensando nessa revisão de conceitos, a gente usa bastante um termo em genética chamado de **“locus”,** ou o plural **“loci”**. O que é um locus? O **locus é a posição física que um gene ocupa dentro de um cromossomo**. É como se a gente estivesse falando do endereço físico no qual esse gene se encontra dentro de um cromossomo, isso é o que a gente chama de locus genômico ou locus cromossômico.

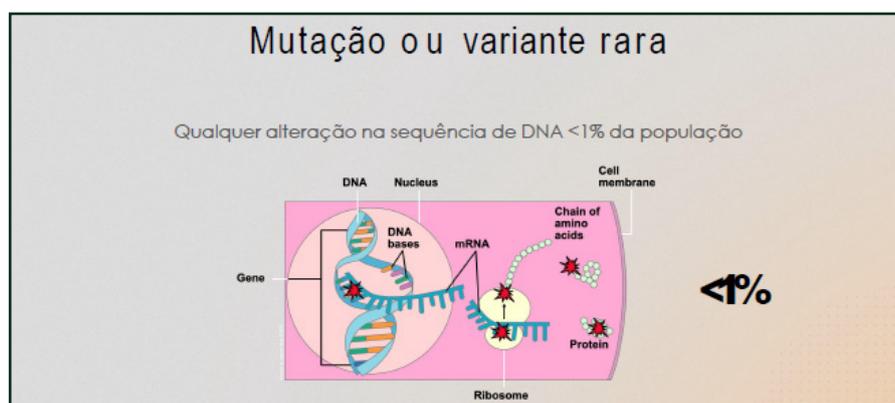


Dentro de um lócus, de um mesmo lócus, a gente pode ter algumas **variações na sequência para esse mesmo lócus**, algumas vezes a gente tem, em cada um dos cromossomos, uma variante diferente, aqui nesse caso, a gente tem uma variante chamada de d (dezinho) e uma variante chamada de D (deção). Com isso a gente tem essa representação em heterozigose, então um sujeito que possui nesse mesmo lócus uma variante D e uma variante d. Ou ainda, esse indivíduo **pode ter duas variantes com o mesmo alelo**, ele poderia ser homozigoto para esse alelo.

### Variabilidade Genética Humana

- A diversidade está baseada na pequena fração da sequência de DNA diferente entre indivíduos
- Genoma Humano
- Similaridade entre os indivíduos: 99,5%
- Diferença entre pessoas: 0,5%

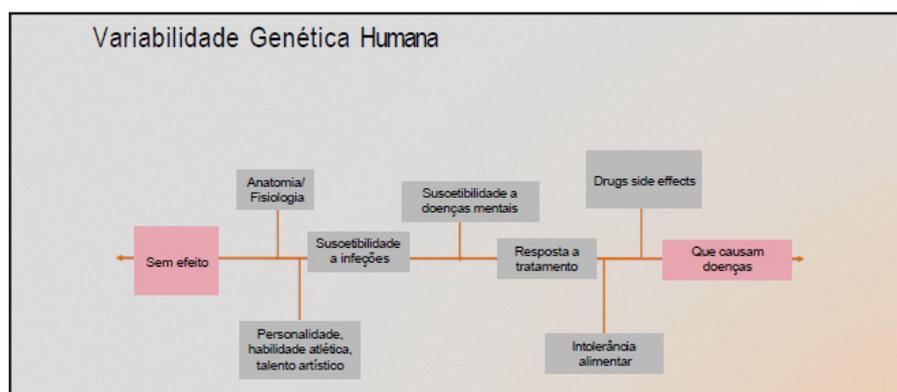
Ainda pensando nessas sequências gênicas que são variantes, a gente pode dizer que essas variantes são responsáveis pela diversidade genética de um indivíduo dentro daquela espécie. Então, se a gente pudesse comparar o genoma, o DNA de duas pessoas da espécie humana, a gente observaria - isso a gente pode fazer - a gente observa uma similaridade de sequência de mais ou menos 99,5%. **Então, dois DNA's diferem em apenas 0,5% da sua sequência**, essa diferença de 0,5% é o que faz com que **as pessoas sejam diferentes**, tanto geneticamente quanto fenotipicamente. E isso garante a diversidade genética de uma espécie.



Essas variantes, essas sequências que são diferentes entre as pessoas, elas podem ser variações que nós chamamos de **variações raras** ou **variações comuns**. As **variantes raras são chamadas também de mutação**. O que é, por definição, uma mutação? Por definição, uma mutação é uma variação na sequência de DNA que está presente em menos de 1% da população. Nesse caso, qualquer variante que esteja presente em menos de 1% de uma população qualquer, a gente pode dizer que essa variante é uma mutação.



Por outro lado, quando essa variante é uma variante que tem uma frequência um pouco maior na população, nós chamamos essa variante de **variante comum ou polimorfismo**. O que são os polimorfismos? Os polimorfismos são variações presentes em mais do que 1% do genoma, lembrando que aqui a gente pode ter diversas variantes, vários polimorfismos, o importante é que cada uma dessas variantes tenha mais do que 1% de frequência na população para ser considerada um polimorfismo. Um dos polimorfismos mais bem conhecidos, mais bem estudados, é o que a gente chama de SNP: *Single-Nucleotide Polymorphism*. Essa variante é **uma variação de um único par de base**, na qual uma das suas variantes, a menos frequente, tem, pelo menos, 1% de frequência na população. Nesse caso, a gente chama essa variante de polimorfismo. E como ela ocupa uma única variante, um único par de bases dentro do genoma, nós chamamos isso de variante do tipo SNP.



Eu representei aqui essas duas variantes, as variantes comuns com o SNP e as variantes raras, mutação, de uma forma geral, mas nós vamos falar um pouquinho mais para frente de cada uma delas. Qual é o efeito dessas variantes? Essas variantes podem ter diversos efeitos na população. O primeiro efeito, a primeira nuance dessa variante, pode ser até uma variante que não causa nenhum efeito na população. Ela só **garante uma variabilidade genética sem efeito funcional**.

Por outro lado, algumas variantes podem simplesmente proporcionar diferenças anatômicas ou fisiológicas para aquele sujeito. Ou até mesmo, permitir uma diferença de personalidade ou de habilidades, habilidade atlética, ou para um talento artístico, ou até mesmo **essa variante pode aumentar o risco de que esse indivíduo tenha uma suscetibilidade a infecções**, por exemplo, a gente viu isso durante a pandemia com a infecção por Coronavírus, **as pessoas têm suscetibilidades diferentes aos mesmos agentes infecciosos**. Ainda, a gente pode ter suscetibilidades diferentes a doenças mentais, diferentes respostas ao tratamento, diferentes intolerâncias alimentares, efeitos adversos de medicação, ou ainda **essas variantes podem causar, de fato, uma doença**.

Então vocês podem perceber que essas variantes, seja uma variante rara ou uma variante comum, elas estão dentro de um espectro de variantes com múltiplos efeitos, cada uma delas tem um efeito que pode ser desde nenhum efeito ou até mesmo causar uma doença. E a gente vai falar um pouquinho mais para frente sobre essa relação do efeito e da frequência da variante.

## PSIQUIATRIA ATUAL

Agora que a gente já fez uma revisão sobre os principais conceitos de genética, falando sobre qual é a singularidade genética entre os sujeitos, o que é uma variante, uma variante comum, uma variante rara; agora a gente vai entrar no contexto da Psiquiatria, agora a gente já vai começar a entrar na genética psiquiátrica de fato.

A Psiquiatria, de forma geral, a Psiquiatria atual, é uma área da Medicina cuja etiologia das doenças ainda é pouco conhecida, cujo diagnóstico ainda é feito por observação do comportamento, não há testes diagnósticos para apoiar o diagnóstico ainda. A detecção, por causa disso, é tardia, porque como o diagnóstico é feito com base na observação do comportamento, muitas vezes você consegue detectar um transtorno mental muito tardiamente.

**A prevenção e a predição das doenças mentais ainda são mínimas na área, e o tratamento ainda é feito por tentativa e erro.** Esse seria o panorama da Psiquiatria atual.

### Psiquiatria atual

- Nenhum teste laboratorial e biomarcadores disponíveis
- Não há evidências genéticas e nem neurobiológicas que dêem suporte para a classificação atual de doenças



Então a gente não tem nenhum teste laboratorial, nem biomarcadores disponíveis no momento, que possam nos ajudar na classificação das doenças ou no diagnóstico ou apoiar essa classificação atual das doenças. Isso ainda não existe. **Esse diagnóstico é feito puramente com base no comportamento**, como a gente falou.

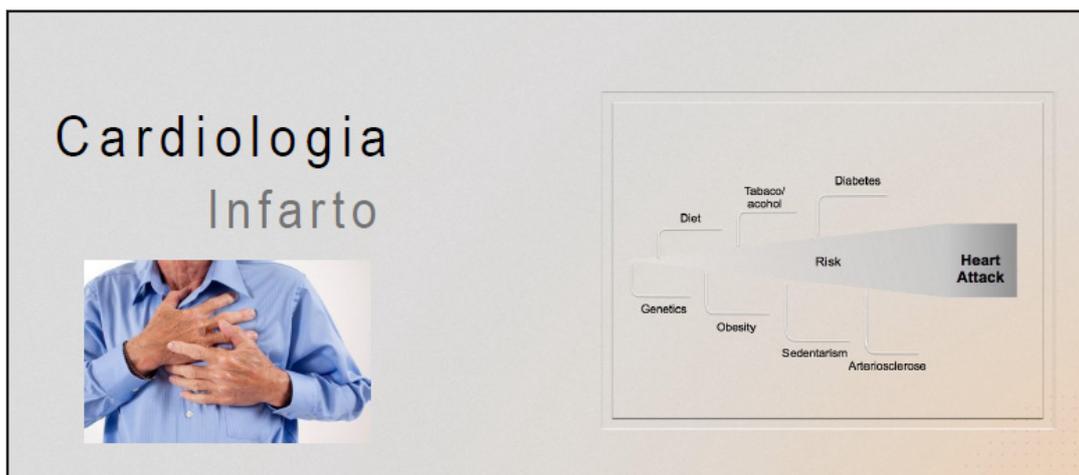
## PSIQUIATRIA DO FUTURO

A Psiquiatria do futuro é uma Psiquiatria em que a gente deseja que seja feito como em muitas outras áreas da Medicina, a gente tem algumas áreas da Medicina em que o diagnóstico já é feito por testes laboratoriais, exames que apoiem esse diagnóstico.

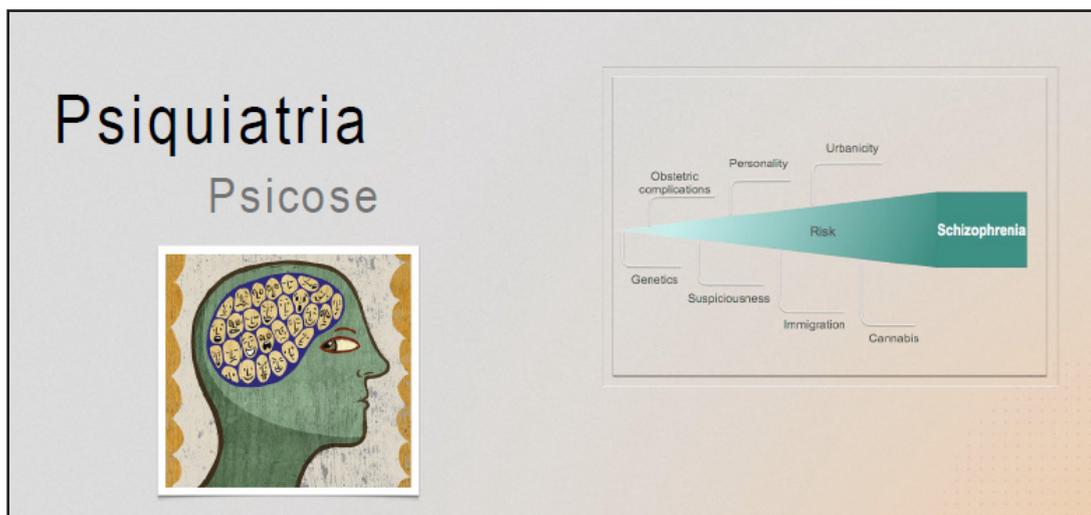


Mas a gente quer para a Psiquiatria do futuro que a compreensão das relações genéticas, da arquitetura genômica, da interação entre os genes encontrados e relacionados às doenças mentais, que isso possa nos ajudar a **entender mais o compartilhamento genético das doenças mentais** e, com isso, talvez, no futuro, a gente possa construir uma classificação mais moderna para esses distúrbios psiquiátricos.

Isso já é feito em outras áreas na Medicina, como na Cardiologia, por exemplo, já se faz, em nível populacional, já existem medidas populacionais de prevenção, na Cardiologia você conhece já os fatores genéticos de risco, conhece os fatores ambientais, sabe que a dieta impacta nesse aumento da suscetibilidade ou na proteção das doenças coronárias. Por exemplo obesidade, uso do álcool, tabaco, sedentarismo. E essa exposição a fatores ambientais pode, então, ir aumentando o risco com o passar do tempo para esses sujeitos, até que o indivíduo culmine com um infarto propriamente dito. Mas você **consegue prever esse infarto e você consegue promover medidas de saúde pública para tentar minimizar as doenças cardíacas de uma forma geral**.



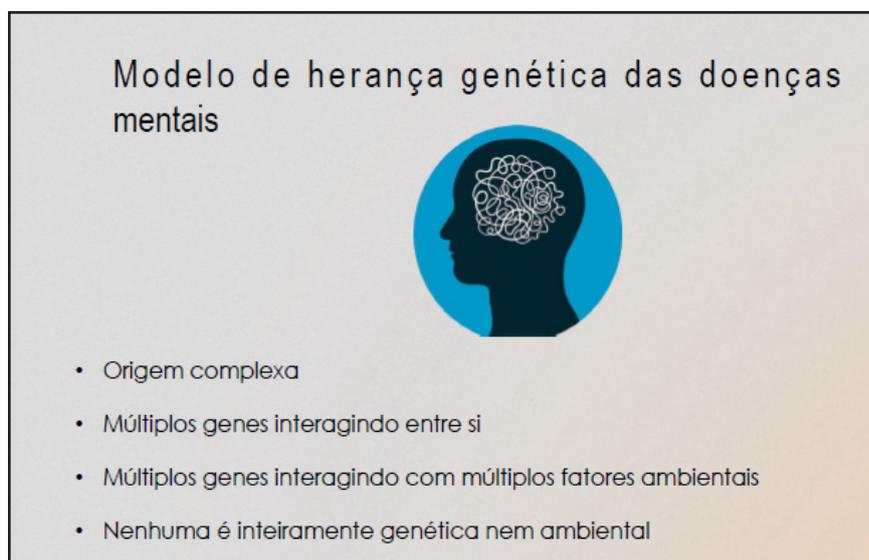
Então a ideia é que a Psiquiatria, no futuro, a gente consiga também fazer esse mesmo tipo de prevenção primária ou secundária ou até mesmo terciária, se a gente conhecer bem os fatores genéticos de predisposição, assim como os fatores de risco ambientais. A gente observa aqui nas Psicoses, de uma forma geral, alguns fatores de risco ambientais que já são conhecidos, como **complicações obstétricas, imigração, urbanicidade, uso de cannabis**. E se a gente conhecer bem esses fatores e como eles estão interagindo com os fatores genéticos, talvez a gente possa, no futuro, fazer políticas, estabelecer políticas de prevenção, como a gente estabelece, por exemplo, em outras áreas da Medicina.



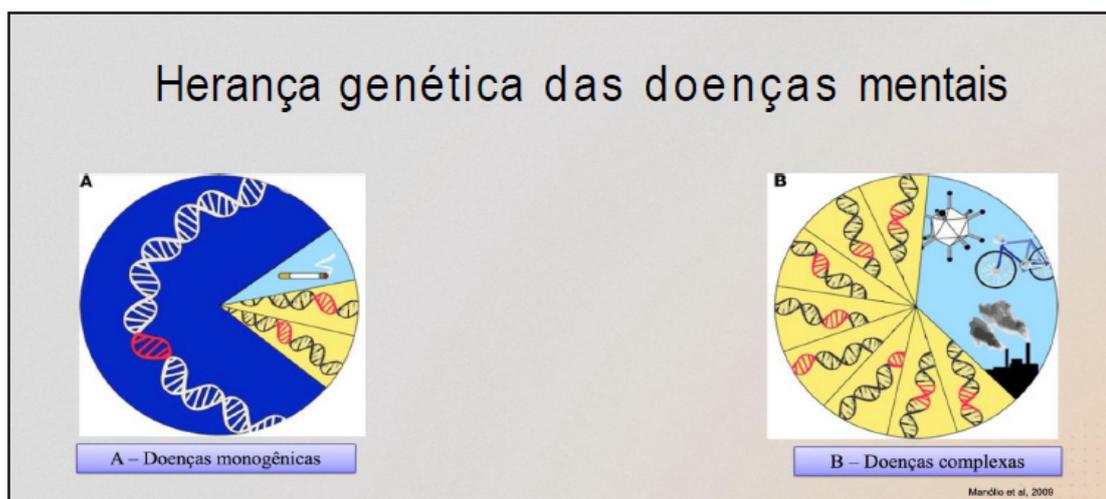
Pensando, então, sobre as bases genéticas desses transtornos mentais, a gente pode dizer que **essas doenças são de origem bastante complexa**, são classificadas como doenças complexas, multifatoriais, nas quais **múltiplos genes interagem entre si simultaneamente e múltiplos genes ainda interagem com fatores ambientais**. Nenhuma delas é inteiramente genética e nenhuma delas é inteiramente ambiental.

A gente pode observar aqui nesse gráfico uma comparação entre as doenças monogênicas e as doenças complexas. As **doenças monogênicas são aquelas doenças genéticas que são causadas por um único par de genes**, por variantes genéticas únicas. Doenças como Acondroplasia, a Hemofilia, são doenças nas quais uma única variante - e no caso são variantes raras - essas variantes são **suficientes para desencadear o fenótipo**. Outras variantes genéticas podem até interagir de alguma forma com a variante causal, mas elas não são necessárias para o desencadeamento do fenótipo.

Também pode haver uma **modulação por fatores ambientais**, mas a variante genética por si só, nas doenças monogênicas, é suficiente para desencadear doenças.

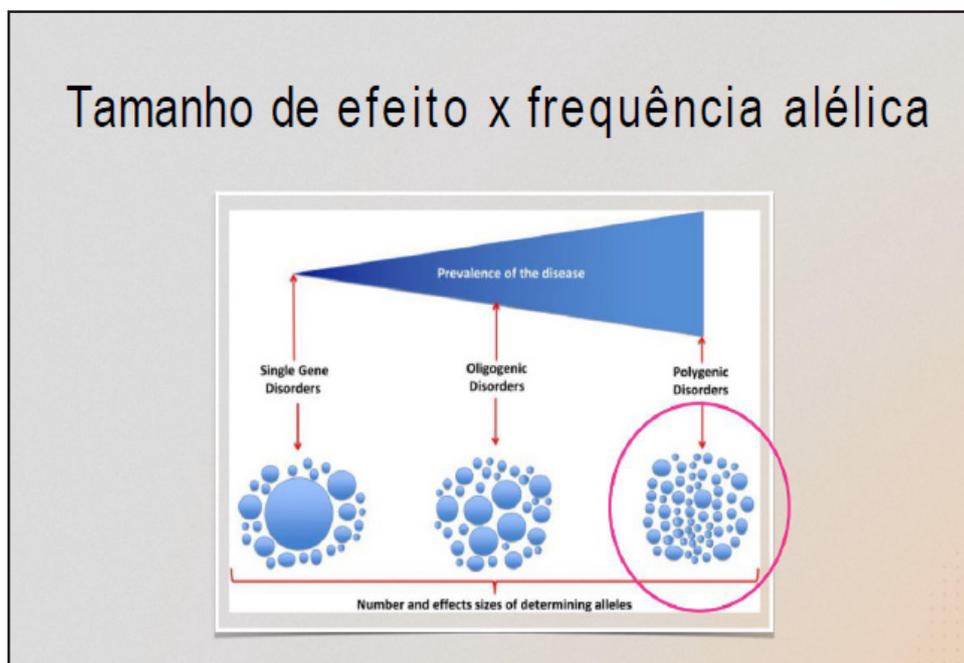


Por outro lado, nas **doenças complexas**, nós não temos nenhuma variante que, sozinha, seja suficiente para desencadear esse fenótipo. Então a gente observa a **participação de diversos genes contribuindo cada um com o seu tamanho de efeito**. Normalmente esses tamanhos de efeito são tamanhos de efeito baixos e o que importa, na realidade, é o conjunto de variantes genéticas que você herda. Essas variantes genéticas interagem com uma grande parcela de fatores de risco ambientais. Por isso que elas são classificadas como doenças complexas; **os transtornos mentais são classificados como doenças complexas** de grande interação gênica e interação gene-ambiente também. Lembrando que aqui as variantes genéticas são variantes genéticas comuns, não são raras, na maioria das vezes, e essas variantes comuns têm baixo tamanho de efeito.



Esse outro gráfico mostra, ele representa, mais ou menos, a mesma ideia, uma comparação, aqui num extremo, as doenças monogênicas, nas quais uma única variante é suficiente. Então vocês podem observar essas bolinhas, essas bolinhas são as variantes gênicas, e o tamanho dessas bolas faz uma **analogia com o tamanho de efeito dessas variantes**.

Você pode observar que essas variantes gênicas das doenças monogênicas têm um tamanho de efeito muito grande. Ao passo que nas doenças poligênicas, multifatoriais, essas variantes têm um pequeno tamanho de efeito, são variantes pequenas, cada uma com o seu efeito.

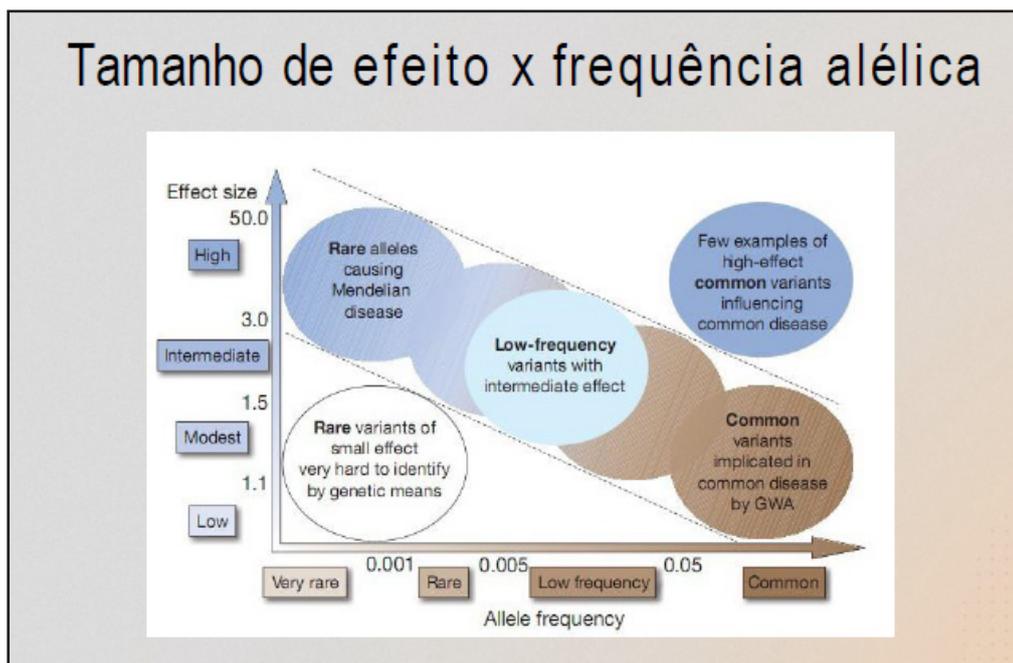


A gente observa aqui nesse gráfico, numa extremidade, no eixo Y, a gente tem o tamanho do efeito dessas variantes, no eixo X a gente tem a frequência dessas variantes, então o que a gente quer mostrar aqui? A gente quer mostrar que nas doenças comuns essas variantes, as variantes associadas às doenças comuns, também são variantes comuns. E, na linha azul, a gente tem que **as doenças mendelianas ou doenças raras, no geral, são causadas por variantes raras**. Certo?

Isso nos diz que variantes comuns são comuns justamente porque o efeito delas é um efeito baixo, porque se elas tivessem um alto tamanho de efeito, muito provavelmente, elas não teriam uma alta frequência na população. Porque o efeito seria tão grande, seria tão deletério, para esse sujeito que porta, é portador dessa variante, que essa variante estaria sofrendo uma pressão seletiva do meio, então ela deixa poucos descendentes, as pessoas se casam menos e, quando se casam, podem ter uma pressão seletiva na qual elas não deixam descendentes.

Então por isso as variantes comuns são comuns e as variantes raras são mais raras, as variantes raras com tamanho de efeito tão grande que são deletérias e não deixam tantos descendentes.

Ainda falando aqui sobre a etiologia desses transtornos mentais, aqui uma analogia que eu faço para mostrar **a interação de gene e ambiente nessas doenças complexas**, então esse jarro pode fazer uma analogia com um sujeito, e esse sujeito nasce com algumas variantes, que são as variantes genéticas que ele herdou dos seus genitores. Essas variantes genéticas estão aqui representadas por bolas, cada uma com um tamanho de efeito, que é o tamanho dessas bolas. Pode ser que um sujeito já nasça com esse jarro com uma grande quantidade de fatores genéticos de risco. Quanto mais fatores genéticos de risco esse sujeito é portador, e quanto maior é o tamanho de efeito dessas variantes, essa carga genética vai ocupar um espaço maior dentro desse jarro.

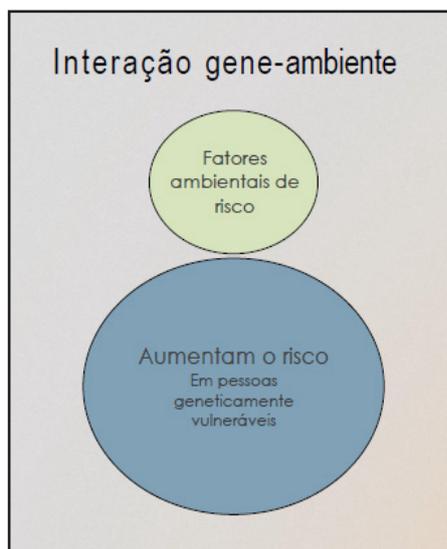


Então é assim que a gente já **nasce: com uma alta suscetibilidade a doenças ou uma baixa suscetibilidade a doenças complexas**. Ao longo da vida, esse sujeito vai sendo exposto a fatores ambientais de risco, que aqui estão representados por estrelas. E que também têm diferentes tamanhos de efeito. Então para cada doença a gente sabe que tem fatores ambientais que têm um tamanho de efeito maior ou menor. E o fato é que esses sujeitos estão expostos a fatores ambientais desde o ambiente intrauterino, e isso **ao longo da infância, adolescência, até a vida adulta; nós somos expostos a esses fatores**. É como se a gente fosse adicionando dentro do nosso jarro fatores ambientais de risco. E quanto mais fatores ambientais de risco a gente tiver ali dentro do nosso jarro, pode ser que esse jarro transborde, e com isso, tenha o aparecimento desses transtornos mentais. **Aparecer ou não aparecer o transtorno mental depende muito da sua carga genética**, da quantidade de fatores de risco que a gente herda, e da quantidade de fatores ambientais a que a gente está exposto.

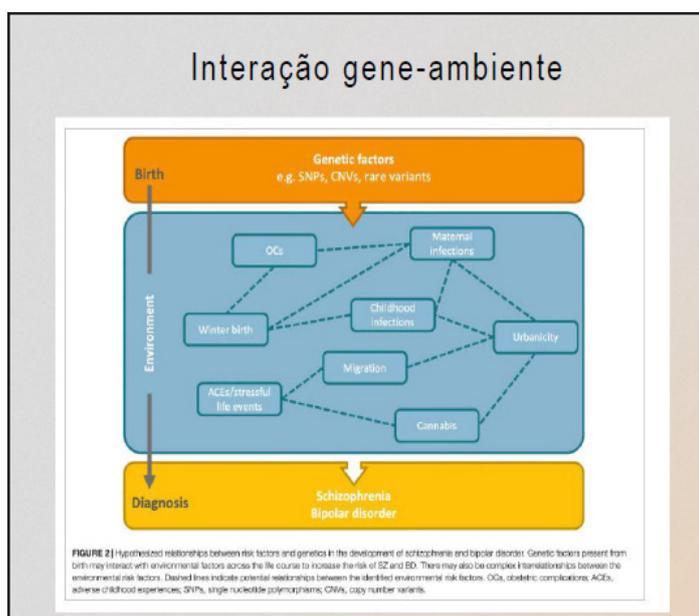
Pode ser que, para um sujeito, ele já nasça com uma grande quantidade de fatores genéticos de risco, mas que ele tenha baixa exposição ambiental de risco. Então, com isso, ele pode até não ter essa determinada doença. Mas, por outro lado, pessoas que até mesmo tenham uma baixa suscetibilidade genética, **quando muito expostas a fatores ambientais, isso pode culminar com a doença**.



O problema da interação gene-ambiente é que as pessoas que já são geneticamente vulneráveis, ou seja, as pessoas que nascem com uma alta suscetibilidade genética, o problema é que **esses fatores ambientais às quais elas estão expostas, vão aumentar o risco dessas pessoas**. Então a gente tem um aumento de risco em pessoas que já são geneticamente vulneráveis. Esse seria o pior cenário; uma pessoa que já tem uma vulnerabilidade genética para esse transtorno e ainda ser exposto a fatores ambientais de risco.

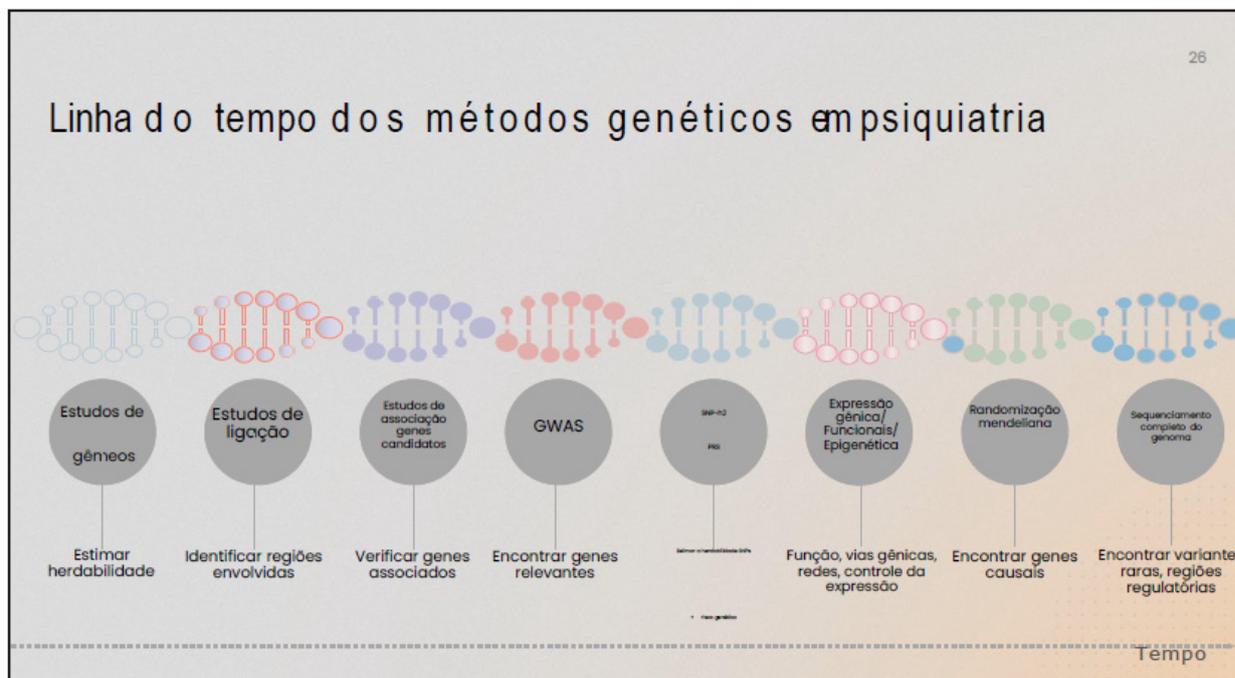


Esses fatores genéticos, a gente já nasce com eles. O que são esses fatores genéticos? São essas variantes genéticas que a gente comentou no começo da aula, que podem ser variantes raras ou comuns. A gente pode observar na literatura, nos livros, nos artigos, que o que importa é o conjunto de variantes. E aí essas variantes são de vários tipos, a gente comentou sobre os SNPs, mas a gente poderia estar falando também de outras categorias de variantes que são também bastante estudadas, como as CNVs, que são as *Copy Number Variations*, ou até mesmo o nosso conjunto de mutações, que a gente apresenta. Tudo isso é uma base para que a gente tenha uma interação com o ambiente no qual a gente está exposto, e isso pode culminar ou não com a doença em questão.



Agora falando sobre métodos, formas para se estudar essa base genética dos transtornos mentais, como que a gente chegou a essas conclusões em relação à etiologia, a interação gene-ambiente, com base nos métodos de estudo que foram desenvolvidos ao longo do tempo.

## EVOLUÇÃO DOS MÉTODOS DE ESTUDO



Aqui eu coloquei para vocês uma linha do tempo mostrando a evolução dos métodos de estudos genéticos que foram aplicados em Psiquiatria. A gente não vai falar, a nossa ideia não é falar sobre todos os tipos de estudos, mas sim falar sobre os principais e os mais utilizados atualmente; então eu vou falar um pouquinho aqui sobre cada um deles e depois eu vou me aprofundar em alguns deles.

Os primeiros estudos que surgiram foram os **Estudos de Gêmeos**. E a ideia nesses estudos de gêmeos era compreender o quanto de uma determinada doença, de uma afecção, é genética. Para isso a gente estudava gêmeos, pares de gêmeos, tanto monozigóticos quanto dizigóticos, e comparando esses gêmeos, a gente conseguia estimar a herdabilidade dessas doenças. A gente já vai falar um pouquinho sobre isso.

Depois vieram estudos que nós chamamos de Estudos de Ligação. Esses estudos de ligação eram formas de **mapear regiões dentro do genoma** que estivessem associadas ou envolvidas, de alguma forma, com as doenças.

Um pouco mais para frente vieram os Estudos de Associação de Genes Candidatos. Nós também vamos falar um pouquinho mais sobre eles, mas a ideia aqui era testar se um gene, de fato, estava associado à doença ou não.

Um pouco mais para frente veio uma era que nós chamamos de Era dos Estudos de GWAS, que são os estudos de *Genome-Wide Association*, a gente usa essa sigla, GWAS, esses estudos foram muito importantes para se encontrar, de uma forma bem exploratória, vários genes envolvidos na doença simultaneamente.

E depois dessa era, que é mais ou menos a era em que a gente tem vivido aqui, agora, no momento, vieram possibilidades de se estimar e fazer cálculos que nos ajudaram um pouco mais com a compreensão genética desses transtornos. Então a gente pôde aprender a fazer novos cálculos para herdabilidade, a gente aprendeu a gerar um **score de risco genético individual** para as pessoas da população.

Depois vieram também estudos um pouco mais funcionais, com base nos resultados do GWAS foi possível, o GWAS foi muito utilizado para dar **pistas biológicas**. E aí, com base nessas pistas, foram feitos estudos mais funcionais, que são os estudos de expressão, os estudos de epigenética, os estudos de vias, para que a gente entenda um pouco mais sobre a função dos genes envolvidos.

Mais recentemente também surgiram os **Estudos de Randomização Mendeliana**, que são estudos genéticos para encontrar, dentro daquelas pistas de genes, encontrar genes que de fato, estão relacionados com a causa da doença.

E ainda, recentemente, apareceram também os sequenciamentos completos do genoma, que nos permitem identificar as variantes raras e não só as variantes comuns. Certo?

## HERDABILIDADE

Bom, falando um pouquinho sobre a herdabilidade, que eu falei para vocês que foi calculada com base nos Estudos de Gêmeos.

### herdabilidade



$h^2 = 2(\text{Taxa de concordância MZ} - \text{Taxa concordância DZ})$

Doença	Concordância MZ	Concordância DZ
Transtorno Bipolar	67%	5%
Epilepsia	70%	6%
Alzheimer	78%	39%
Autismo	90%	4.5%

Essa herdabilidade é um índice representado pelo símbolo  $H^2$ , e esse índice é um índice que varia de 0 a 1, ou de 0% a 100%, se estivermos falando em porcentagem, no qual ele **está estimando a taxa de contribuição genética para um determinado transtorno**, para uma determinada característica ou doença. Esse índice foi calculado, como eu mencionei para vocês, com base em estudos de gêmeos, então o que se fazia era comparar pares de gêmeos monozigóticos e pares de gêmeos dizigóticos. E estimar a taxa de concordância do fenótipo entre esses pares; qual porcentagem de pares, entre todos os estudados, em que os dois gêmeos apresentavam esse fenótipo, essa era a taxa de concordância do fenótipo e isso a gente compara entre gêmeos monozigóticos e dizigóticos. Se você tem uma taxa de concordância maior para gêmeos monozigóticos, isso quer dizer que a contribuição genética é maior do que a contribuição ambiental.

A gente observa aqui nesse quadro alguns transtornos, no qual a gente observa a contribuição genética para esses transtornos, aqui a gente tem um quadro mostrando a **taxa de concordância entre gêmeos monozigóticos e a taxa de concordância entre gêmeos dizigóticos**. Então, quanto maior é essa diferença entre gêmeos monozigóticos e dizigóticos, significa que a gente tem uma contribuição maior genética do que ambiental.



Com base nessa fórmula que eu mostrei para vocês, embora existam outras fórmulas para se calcular herdabilidade baseada em gêmeos, com base nisso, a gente pode observar aqui um quadro mostrando alguns exemplos de herdabilidade entre os distúrbios psiquiátricos.

Então a gente observa aqui que entre os mais altos estão síndromes mais relacionadas ou doenças mais relacionadas à contribuição genética. Por exemplo, o Transtorno Bipolar, Esquizofrenia, e a gente tem doenças que ainda não estão aqui nesse quadro, mas que têm uma contribuição genética menor, como a Depressão, por exemplo. Então a gente consegue, com base nesse índice, estimar o nível de contribuição genética para esses transtornos.

## ESTUDOS DE ASSOCIAÇÃO

Como eu falei para vocês, eu vou dar uma pincelada aqui sobre alguns desses métodos de estudos que nos levaram a compreender a etiologia, ou pelo menos a ter alguma noção de etiologia, entre os transtornos mentais. A gente já falou aqui sobre herdabilidade baseada em gêmeos, mostrando que esses estudos de gêmeos foram muito importantes lá, algumas décadas atrás, para nos mostrar que a contribuição genética para essa doença, qual a contribuição genética para essa doença. E depois vieram os estudos de associação.

Os **estudos de associação de genes candidatos** foram estudos bastante importantes, só que como você tem que ter gene candidato, você precisa ter um gene que você diga que esse gene, você tem uma pista sobre ele, você sabe, com base em outros estudos, com base nos estudos de ligação que vieram antes, você tem uma pista de que aquele gene pode ou não estar envolvido com uma doença.

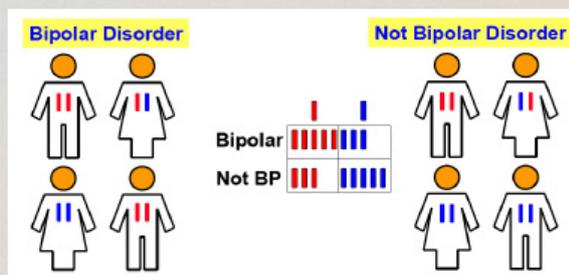
## Estudos de associação de genes candidatos

- Tem que ter hipótese "a priori"
- Escolher gene(s) plausíveis
- Identificar variante funcional
- Problems:
  - N small
  - What gene to study? What variants?

Então, para realização desse tipo de estudo a gente precisa ter uma hipótese *a priori*, uma **hipótese biológica a priori**. E com isso a gente escolhe, tem que escolher alguns genes para se estudar e **poder testar se esse gene está associado ou não com a doença**. Alguns problemas desse tipo de estudo é que os estudos que foram realizados tinham N amostral muito pequeno, e também sempre tinha a dúvida de qual gene estudar, que gene a gente vai testar, se ele de fato tem alguma pista biológica anterior, você tinha que tomar essa **decisão da escolha dos genes**.

## Estudos de associação de genes candidatos

- Co-ocorrência (estatisticamente significativa) de determinados alelos e fenótipos
- Pode ser uma associação biológica ou estatística
- Por seleção natural
- Por estratificação populacional



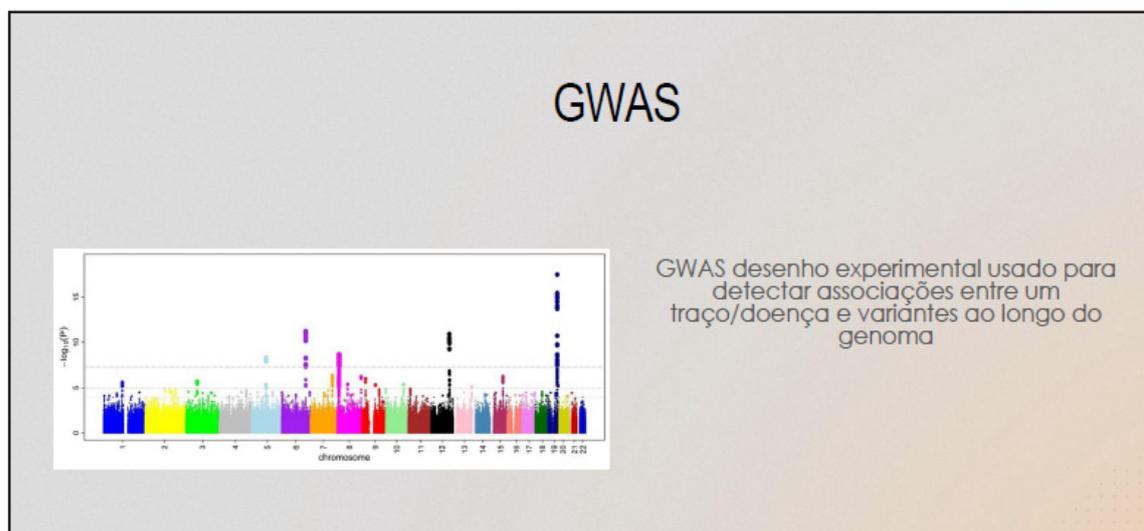
Esse tipo de estudo se baseia tanto em **estudos de comparação entre grupos de caso e controle** como estudos de **comparação entre famílias**. O mais comum é estudar caso e controle. E o que a gente pretende nesse tipo de estudo? A gente pretende testar a hipótese de que aquele gene está associado ou não com a doença. O que a gente vai fazer? A gente vai genotipar esses sujeitos dos dois grupos, caso e controle, para **determinar quais são as variantes genéticas que esses dois grupos possuem**. Com base nisso, a gente vai fazer um teste estatístico para verificar se a **frequência daqueles alelos** é maior em um grupo do que em outro, e se essa diferença é estatisticamente significativa ou não. Se a gente encontra uma diferença de frequência entre casos e controles e essa diferença é estatisticamente significativa, a gente diz que esse gene está associado com a doença.

Esse tipo de estudo tem alguns problemas, além dos números amostrais serem baixos, pequenos, no começo, quando esses estudos surgiram, outro problema era você controlar esse tipo de estudo pela ancestralidade genética dessas pessoas. Por quê? Porque se você encontra uma associação de um gene, de uma variante com a doença, pode ser que essa diferença de frequência que você encontrou seja uma diferença não por causa da doença em si, mas sim porque esses **grupos são diferentes geneticamente**. Então, havia um problema metodológico aqui, em que você tinha poucas formas de ajustar esses estudos para a ancestralidade desses grupos.

## GWAS - GENOME WIDE ASSOCIATION STUDY

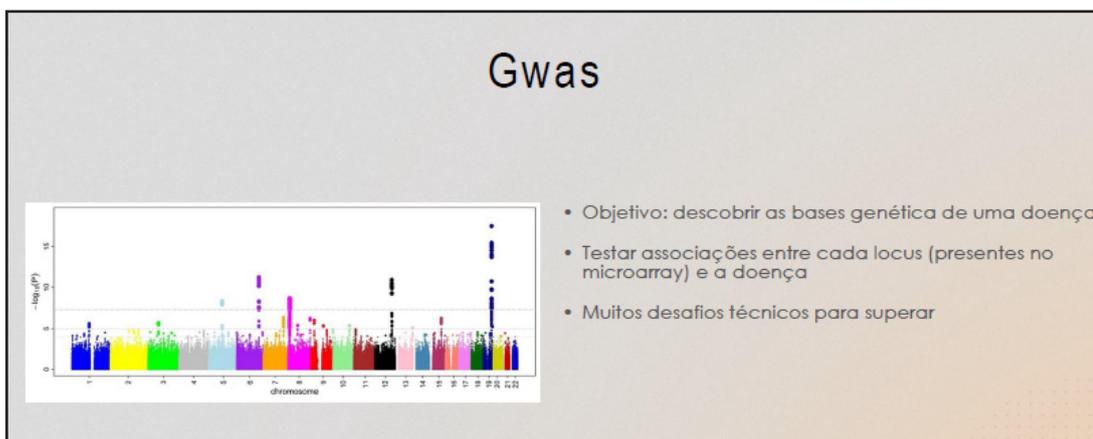
E a gente vai ver depois que, atualmente, a gente vem conseguindo superar essa dificuldade nesse novo tipo de estudos de associação, que são os estudos de associação em larga escala.

Primeiro surgiram os estudos de associação gene candidato e depois, com a evolução das técnicas, a gente teve a oportunidade de estudar milhares, ou até milhões de genes, simultaneamente, com uma **varredura genômica**. E com isso a gente conseguiu fazer associações em larga escala, daí o nome *Genome-Wide Association Study*. São estudos de associação em larga escala, então ao invés de você testar um único gene, sendo o foco uma ou outra variante desse gene, aqui você consegue estudar milhões de variantes simultaneamente; você testa a associação de milhões de variantes simultaneamente.



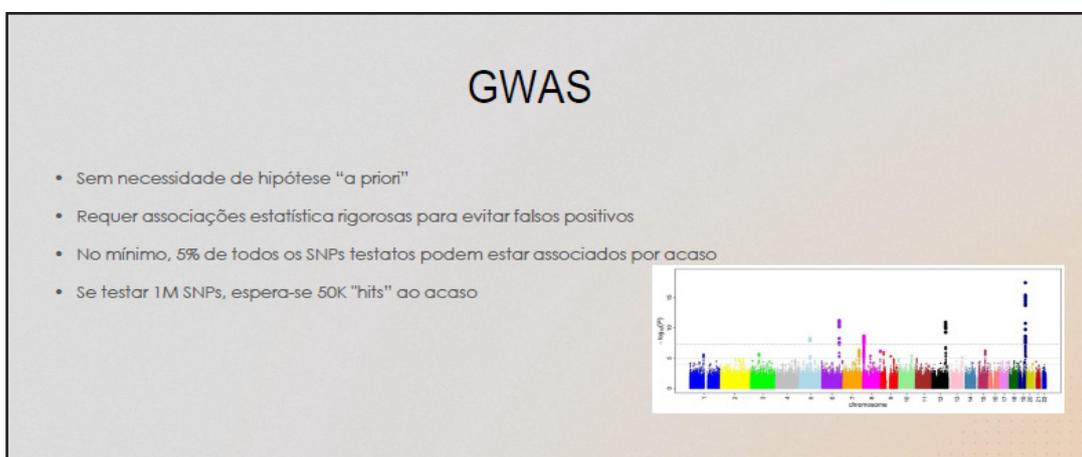
Esse tipo de estudo nos mostra um gráfico como resultado, esse gráfico que vocês estão vendo aqui é o gráfico que é gerado após a realização de um GWAS, a chama-se gráfico de *Manhattan Plot*, porque ele forma picos de associação bem altos, que lembram uma série de prédios de Manhattan, daí o nome *Manhattan Plot*. Essa visualização gráfica vai nos **mostrar se existem loci associados a essa doença em questão**; você faz aqui uma correção estatística para evitar falsos positivos e, depois dessa correção, se essas associações permanecem, como vocês estão vendo nessa representação aqui; a gente tem duas linhas tracejadas, a linha mais alta é o nível de significância estatística, então aquelas regiões genômicas que sobreviveram a essa correção estatística, a gente pode dizer que essas regiões estão associadas com a doença.

Nesse caso, como a gente está testando milhares ou milhões de variantes simultaneamente, nesse caso aqui a gente está testando milhões de SNPs simultaneamente, a gente **consegue ter uma ideia geral sobre a base genética da doença e não testar gene a gene**, aqui a gente consegue ter uma visão mais ampla sobre a contribuição genética para essas doenças. É claro que, como a gente está fazendo um teste estatístico para milhões de associações simultâneas, a gente tem muitos desafios técnicos para superar nesse tipo de método.

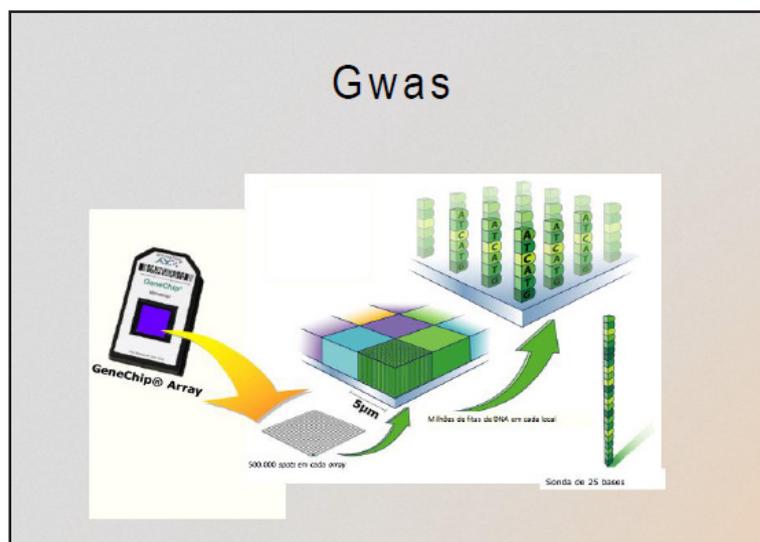


O bom desse método é que você não precisa de uma hipótese *a priori*, como você precisa no gene candidato, aqui você está testando variantes ao longo do genoma, em todos os cromossomos, em todos os cromossomos você está testando associação. Então você não precisa ter uma hipótese, *a priori*, você pode não ter nenhuma ideia sobre a base genética daquela doença e aí você vai utilizar esse tipo de metodologia.

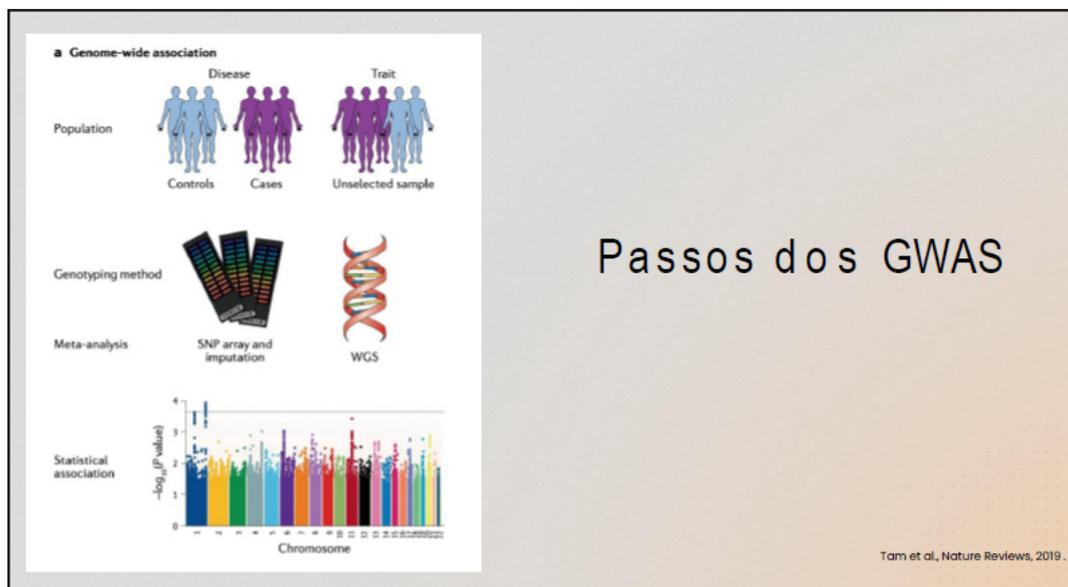
Por outro lado, como você está testando muitas regiões, você precisa ter um **rigor maior para as correções estatísticas**, para que a gente evite os falsos positivos. Então, se a gente está testando, por exemplo, um milhão de SNPs simultaneamente, e a gente adota um índice alfa de 5%, que é o valor de  $P_{ra}$ , isso significa que 5% de todos os SNPs que a gente encontrou como associados podem estar associados por acaso, com uma chance de a gente estar errado ao afirmar que eles estão associados. Então a gente faz uma correção mais rigorosa, para tentar remover esses SNPs que são, na verdade, **falsos positivos**, para evitar essas associações espúrias.



Aqui a gente tem a representação do chip que a gente utiliza dentro do laboratório - esse é um dos modelos e marcas, a gente tem diversos tipos - e a gente pode, com base nessa lâmina, **nesse chip, colocar o DNA de uma pessoa** lá dentro e testar quais são as variantes genéticas que ele tem.



E com isso a gente vai comparar dois grupos, o grupo de caso e o grupo de controle, pessoas que não têm a doença, e fazer essa associação SNP a SNP, gerando esse gráfico que nós chamamos de *Manhattan Plot*.



O *Manhattan Plot* é construído com base em tabela, num *output* mais ou menos como esse que vocês estão vendo, a gente tem elencado em tabelas bastante grandes computacionalmente, a gente tem tabelas com mais de milhões de linhas, e aí a gente observa nessa tabela que a gente tem qual é aquele SNP que foi testado, quais são as variantes genéticas que aquela pessoa tem para aquele SNP, esse beta que nós estamos vendo aqui nessa coluna é um **valor de tamanho de efeito**, é o que a gente chama de *odds ratio* ou, nesse caso, é o tamanho do efeito dessa variante, então com isso, a gente consegue saber qual é o tamanho do efeito dela sobre esse fenótipo. E, também o valor de P do teste estatístico, ou seja, se aquela variante está ou não estatisticamente associada.

## Esse é o resultado do GWAS

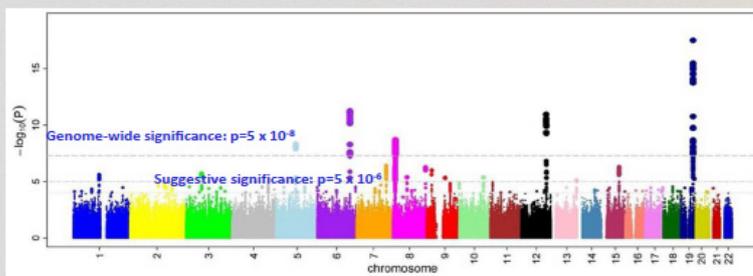
Chr	SNP	bp	Tamanho de efeito (OR, $\beta$ )		Freq	b	SE		p
			A1	A2			se	Qual a variabilidade do $\beta$ ?	
1	rs2286139	761732	C	T	0.1379	-0.0104056	0.00732416	0.155397	
1	rs12562034	768448	A	G	0.10475	-0.00627592	0.00827054	0.447955	
1	rs4970383	838555	A	C	0.247975	0.00946201	0.00587444	0.107243	
1	rs1806509	853954	C	A	0.3912	0.0152744	0.00523012	0.00349507	
1	rs13302982	861808	A	G	0.018025	-0.0180122	0.0189517	0.341895	
1	rs28576697	870645	C	T	0.29355	0.0116486	0.00556379	0.0362916	
1	rs2340582	882803	A	G	0.05465	0.0119371	0.0111055	0.282426	
1	rs3748594	886384	A	G	0.025975	-0.01244	0.0158797	0.433401	
1	rs28504611	908414	T	C	0.022225	0.00388796	0.0171623	0.820781	
1	rs9777939	929190	A	G	0.03345	-0.00446279	0.0141522	0.752502	
1	rs1891910	932457	A	G	0.2295	-0.00647527	0.00605864	0.285175	
1	rs35940137	940203	A	G	0.050575	-0.10689	0.0115935	2.97533e-20	
1	rs6657048	957640	T	C	0.011825	0.0892934	0.0233322	0.000129691	
1	rs9803031	987200	C	T	0.083875	-0.00434284	0.0091306	0.634334	

- Each SNP generates an O.R. (an effect size for that phenotype)

Com base nisso a gente constrói essa visualização gráfica, que é o *Manhattan Plot*, para tentar estudar essas pistas biológicas. A interpretação desse tipo de estudo de GWAS deve ser bastante cautelosa. Por quê? Porque como a gente viu lá no começo da aula, o tamanho do efeito dessas variantes são variantes de baixo tamanho de efeito. Aqui é um gráfico mostrando o *odds ratio*, ou seja, o tamanho de efeito das variantes encontradas, por exemplo, num GWAS de Esquizofrenia, em 2014.

A gente observa que as variantes que têm o maior tamanho de efeito estão aumentando o risco em 20% ou em 30% nos portadores dessas variantes. Então o tamanho de efeito é um tamanho baixo. Por isso a gente deve **ter cautela para interpretar esse tipo de estudo** e mais, a gente tem que utilizar essas regiões genômicas que foram associadas estatisticamente simplesmente como **pistas biológicas** e continuar os estudos a partir dali. A gente não termina os estudos, a partir do GWAS a gente não consegue inferir relação de causalidade ou de funcionalidade dessas variantes genômicas. A gente precisa fazer outros estudos para testar isso, então a gente utiliza isso apenas como pistas.

## Plot view of the results (Manhattan Plot)



The genomic regions associated with the phenotype rise like Manhattan skyscrapers

Esse tipo de método teve alguns problemas metodológicos também, um deles é a **replicação**. Então, vários estudos de GWAS realizados, espalhados pelo nosso planeta, a gente tem que muitas vezes um grupo de estudo encontrou genes relacionados a uma doença, associados a uma doença, e quando um outro grupo, em outro país, com outra população de estudo, foi testar ou replicar isso, não encontrou essa mesma associação.

## Interpretação deve ser cautelosa

- Variantes comuns associadas em abundância, mas com modesto OR
- Investir em estudar estas pistas biológicas

Odds-ratio vs risk allele frequency, PGC-schizophrenia 2014 results  
(N ~ 35,000 cases, n=128 genome-wide significant variants)

Então a **replicação é um grande desafio**. E ela é um grande desafio por vários motivos. O primeiro motivo é o **tamanho da amostra** que essas pessoas utilizam para esse tipo de estudo. E outro desafio é a **questão da ancestralidade genética** que a gente comentou, essas pessoas já são geneticamente diferentes, e isso já gera um viés no estudo.

## Replicação é um desafio



Então, tanto a replicação foi um problema, e essa replicação foi um problema justamente por causa do tamanho da amostra utilizada, como do tipo de população utilizada.

A gente vai agora finalizando essa aula e, para a próxima aula, o que a gente vai abordar? A gente vai abordar qual seria o tamanho ideal para esse estudo, uma vez que eu estou dizendo que o tamanho da amostra é fundamental para você melhorar as descobertas ou aumentar o poder das descobertas genéticas, também me refiro ao tipo de ancestralidade que essas pessoas têm.

Então a gente vai falar sobre o tamanho da amostra, vamos falar sobre ancestralidade e vamos falar também sobre análises pós o GWAS, o que a gente pode encontrar como medidas genéticas que nos auxiliem a compreender melhor os transtornos mentais.

Eu encontro vocês na próxima aula.

Obrigada.

Tchau, tchau.

## AULA 02

Qual o tamanho de amostra ideal para encontrar variantes de baixo efeito?

Olá, pessoal. Bem-vindo à segunda parte da nossa aula. Vamos abordar agora algumas questões a respeito dos estudos que a gente comentou na aula passada, dos estudos de associação em larga escala. E pegando o gancho daquela aula, a gente vai conversar agora sobre qual é o **tamanho da amostra ideal para encontrar variantes de baixo efeito** que são variantes encontradas nos transtornos mentais, relacionadas com os transtornos mentais.

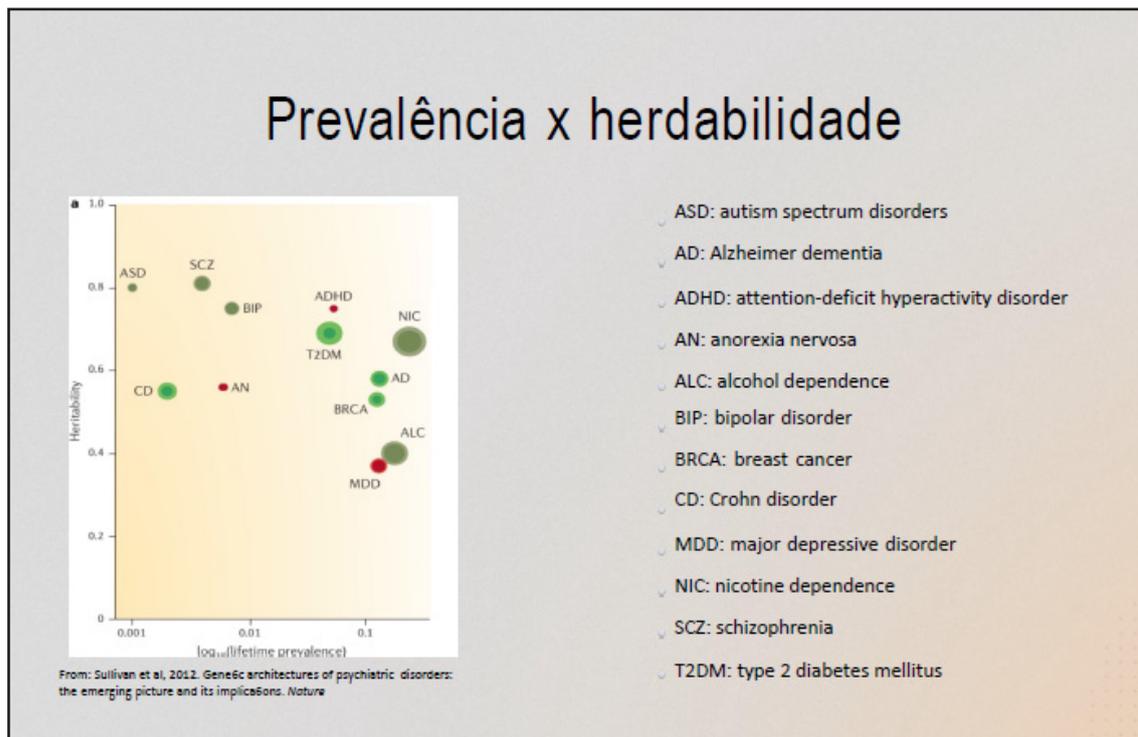
### Tamanho de amostra é crucial

- Para detectar variantes de baixo efeito-> amostras enormes
- Depende da prevalência da doença

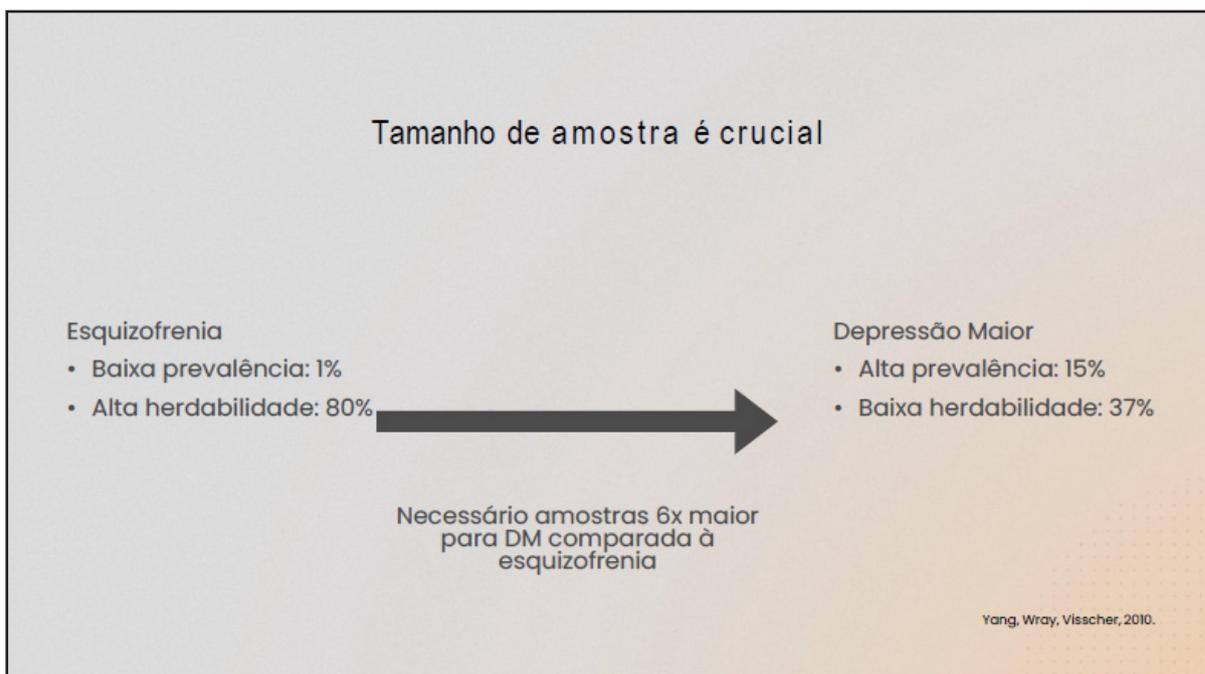


O tamanho da amostra é crucial para esse tipo de estudo; o tamanho da amostra é uma questão bastante importante nesse tipo de estudo e **quanto maior a amostra melhor e maior é o poder de descoberta das variantes**, certo?

Entretanto, isso **depende muito da prevalência da doença**. Se a gente está falando sobre doenças que são muito prevalentes, nós precisamos de tamanhos amostrais maiores. E se a gente está falando de doenças menos prevalentes, a gente precisa de tamanhos e amostras menores. Nesse gráfico, a gente observa aqui em seus extremos, na representação ASD aqui, que as doenças do Transtorno do Espectro Autista, gente observa que essas doenças são doenças mais raras, ou seja, menos prevalentes e com isso a gente precisa de tamanhos de amostras maiores. Assim como Esquizofrenia.



E por quê? Porque essas doenças têm uma **alta herdabilidade**. Ou seja, uma alta contribuição genética. No outro lado da curva, ali em vermelho, temos o, por exemplo, a Depressão Maior, representada pela sigla MDD. A depressão é uma doença bem mais prevalente do que a Esquizofrenia, por exemplo, e com uma herdabilidade mais baixa, ou seja, com uma contribuição genética menor em relação à Esquizofrenia, por exemplo. Com isso a gente tem as doenças que são mais prevalentes, para que a gente encontre as bases genéticas dessas doenças, a gente precisa de amostras muito maiores.



Comparando esses dois extremos, Esquizofrenia e Depressão, por exemplo, a gente sabe, por meio de modelagens, de estudos de modelagem, que a gente precisa de uma amostra seis vezes maior para depressão quando comparada a uma amostra de Esquizofrenia, para poder encontrar essas variantes relacionadas a essas doenças.



## PGC overview



- 2007 - atual
- 800+ pesquisadores, maior consórcio
- 150+ instituições
- 40+ países
- Até fim 2019: 320 papers
- Resultados disponíveis e gratuitos

- Schizophrenia
- ASD
- ADHD
- MDD
- Bipolar disorder
- Eating disorders
- PTSD
- Alzheimer's
- Anxiety disorders
- OCD
- Substance abuse

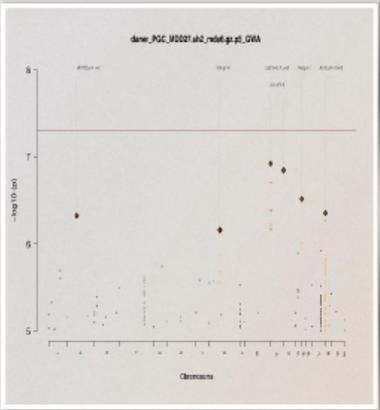
<https://www.medunc.edu/pgc/>

Para garantir esse grande número amostral dos estudos de Genome-Wide Association, alguns consórcios mundiais vêm sendo realizados. O maior consórcio de Genética Psiquiátrica até o momento é o PGC - Psychiatric Genetics Consortium, é um consórcio que teve o seu início em 2007; tem quase mil pesquisadores envolvidos, mais de quarenta países. O Brasil é um deles bastante envolvido nesse consórcio e a ideia é fazer colaborações internacionais para que a gente possa aumentar o poder da descoberta dessas variantes genéticas.

## Tamanho de amostra é crucial



### MDDgwas



data: PGC\_MDD21\_ah2\_results\_gwasc

2013, PGC o maior GWAS com resultados negativos da história

18K casos, 35K controles

Entretanto, a despeito desses esforços mundiais, a gente no começo não conseguia grandes amostras; a gente tinha tamanhos amostrais ainda muito reduzidos. Em 2013, por exemplo, esse consórcio apresentou o maior Gwas com resultados negativos da história. Ou seja, era um Gwas de Depressão, na qual utilizaram 18.000 casos e 35.000 controles, e mesmo com esse tamanho amostra bastante grande não houve **nenhuma associação genômica relacionada à Depressão** neste estudo, o que foi bastante decepcionante.

**Tamanho de amostra é crucial**

**MDD gwas**

**A mega-analysis of genome-wide association studies for major depressive disorder**

Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium

Molecular Psychiatry 18, 401-411 (2013) | Download Citation &

**Abstract**

Prior genome-wide association studies (GWAS) of major depressive disorder (MDD) have met with limited success. We sought to increase statistical power to detect disease loci by conducting a GWAS mega-analysis for MDD. In the MDD discovery phase, we analyzed more than 1.2 million autosomal and X chromosome single-nucleotide polymorphisms (SNPs) in 18 759 independent and unrelated subjects of recent European ancestry (9240 MDD cases and 9519 controls). In the MDD replication phase, we evaluated 354 SNPs in independent samples: 6783 MDD cases and 50 855 controls. We also conducted a cross-disorder meta-analysis using 819 autosomal SNPs with  $P < 0.0001$  for either MDD or the Psychiatric GWAS Consortium bipolar disorder (BIP) mega-analysis (8238 MDD cases/8039 controls and 6998 BIP cases/7775 controls). **No SNPs achieved genome-wide significance.**



Nenhum SNP foi encontrado com o nível de significância pra Gwas. Por que disso? A gente sabe o porquê, porque a **Depressão ela é uma doença bastante prevalente, com uma baixa herdabilidade**. Pra você descobrir as causas genéticas, você precisa de um N muito grande, uma vez que a contribuição genética é menor e a prevalência é maior.

**Tamanho de amostra é crucial**

**MDD Gwas**

**Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression**

Naomi R. Wray, Stephan Ripke, [...] the Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium

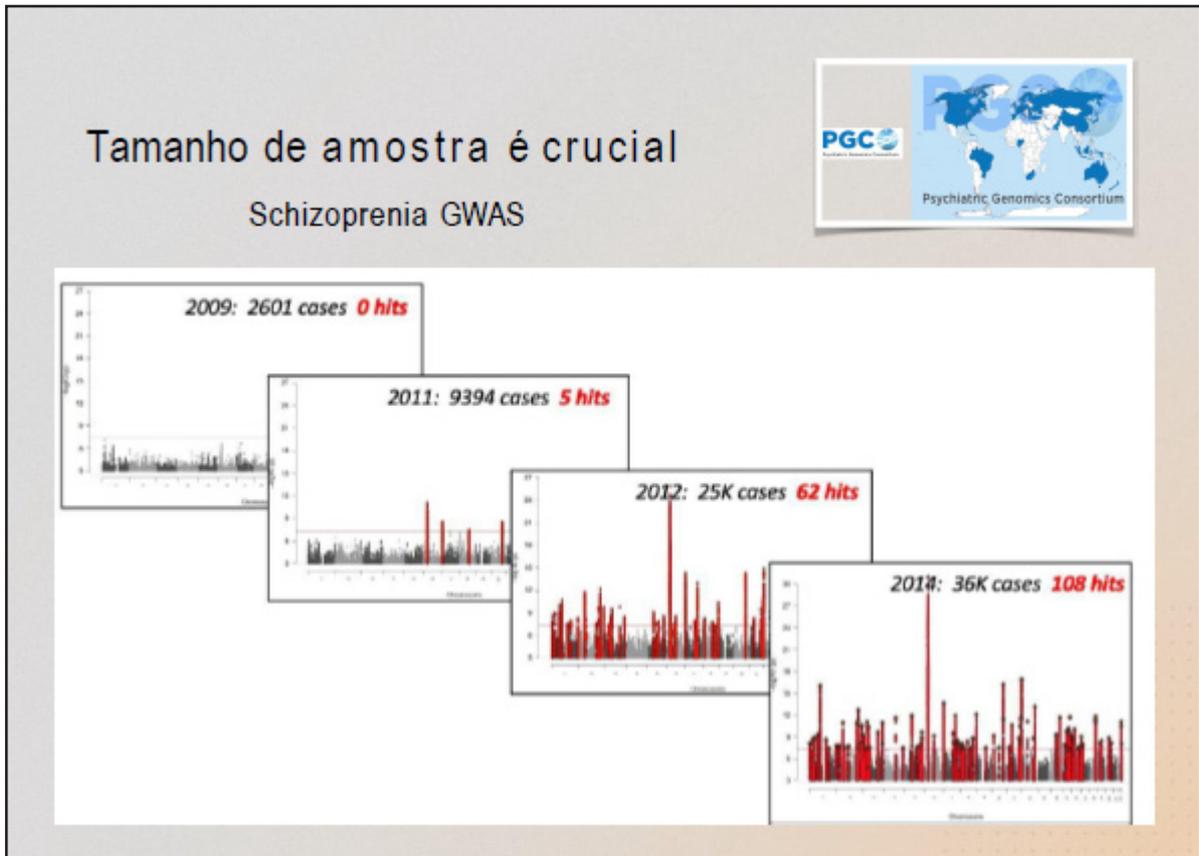
Nature Genetics 10, 668-681 (2018) | Download Citation &

**Abstract**

Major depressive disorder (MDD) is a common illness accompanied by considerable morbidity, mortality, costs, and heightened risk of suicide. **We conducted a genome-wide association meta-analysis based in 135 458 cases and 344 901 controls and identified 44 independent and significant loci.** The genetic findings were associated with clinical features of major depression and implicated brain regions.



O tamanho da amostra é crucial nesse tipo de estudo.



Esse tipo de estudo ele precisa de Ns muito maiores. No segundo Gwas que esse consórcio mostrou, esse mesmo grupo agora com uma amostra bem maior, agora aqui com 350.000 controles e 135.000 casos, nós pudemos encontrar 44 regiões significativas.



Para vocês verem que o tamanho da amostra é muito relacionado ao poder da descoberta; certo? E a mesma coisa aconteceu pra Esquizofrenia. Os primeiros de Gwas relacionados não encontraram nenhum Hit; depois com um pouco mais de 25.000 casos foram encontrados 62 Hits; depois 36.000 casos, 108 Hits. Quando eu me refiro a **Hits**, eu estou dizendo **108 regiões do genoma associadas com a doença**.

## What is the end point of GWAS?



- Projeções sugerem >250K casos são necessários para esquizofrenia e identificar a maioria das variantes.
- >1 milhão para MDD
- Muito menos é necessário para testar correlações genéticas, score poligênico e análise de vias

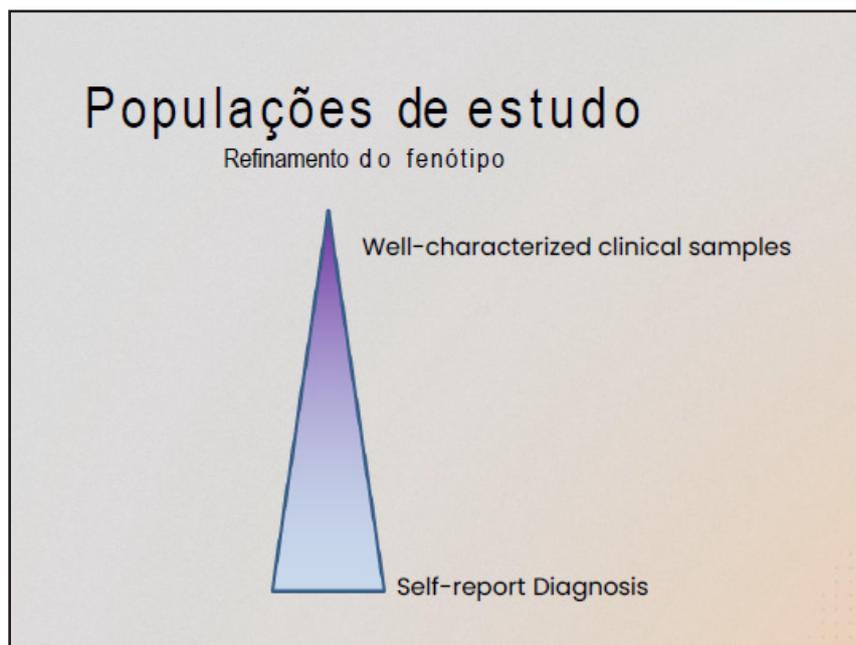
Esse ano, no início do ano, esse papel que foi publicado pela revista Nature desse consórcio mesmo de genética da Esquizofrenia, agora com uma amostra muito maior passou de 108 *loci* encontrados para 270 *loci* encontrados. Isso só mostra que **quanto maior a amostra maior é o poder de você descobrir a base genética de uma doença**, principalmente quando a gente está falando de doenças bastante prevalentes.

Se eu estou dizendo que quanto maior a amostra maior é o poder. Qual seria o fim? Onde está o fim? Dos estudos de Gwas? Bom a gente sabe que uma hora a gente vai atingir um platô; a gente não atingiu ainda esse platô de descoberta, mas a gente ainda está na fase de aumentar exponencialmente, mas a gente sabe que uma hora a gente vai atingir, mas isso ainda vai depender do tamanho da amostra utilizada. A gente sabe que, por exemplo, para depressão, algumas previsões mostram que a gente vai precisar de pelo menos um milhão de casos de depressão para que a gente descubra a grande maioria das regiões associadas a esse tipo de doença.

## ESTUDOS POPULACIONAIS SÃO CRÍTICOS

Os estudos populacionais têm várias questões como, por exemplo, além do tamanho da amostra, o **tipo de fenótipo estudado**.

A gente sabe que quanto melhor o fenótipo, ou seja, mais bem caracterizado clinicamente, também é maior o poder da descoberta; amostras em que as pessoas fazem um, a gente chama de Self Report, ou seja, que elas mesmas dizem que já tiveram diagnóstico, esse tipo de fenótipo é um fenótipo um pouco frouxo, muito diferente de quando você tem um refinamento clínico maior. E com esse refinamento clínico a gente consegue um poder de descoberta maior, amostras mais homogêneas clinicamente e mais bem caracterizadas.



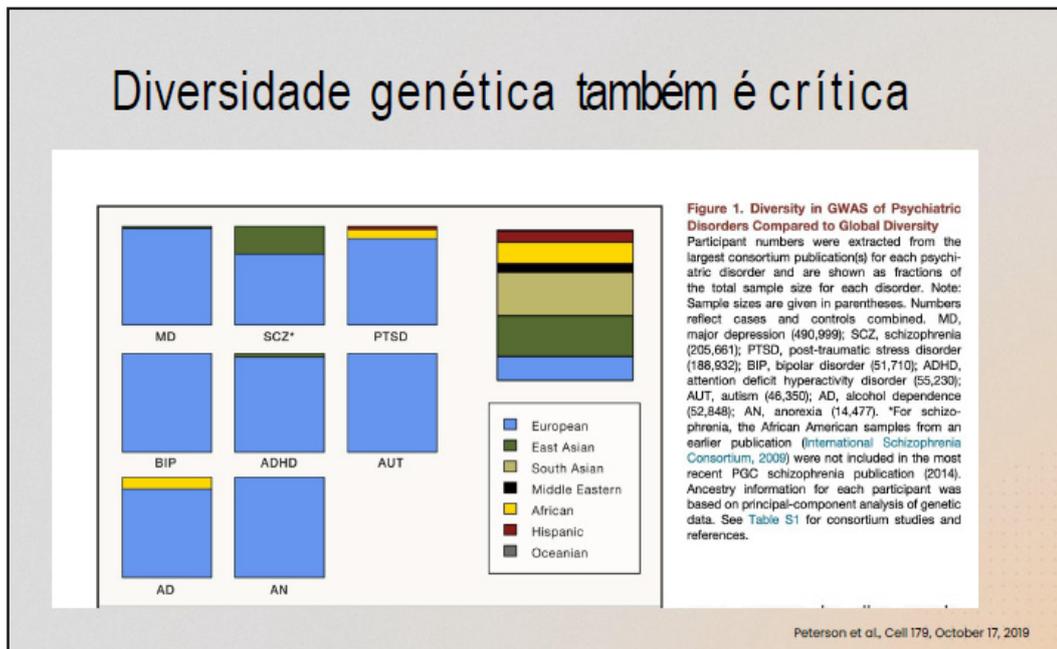
Outro problema desse tipo de estudo como eu comentei na aula passada é a **ancestralidade genética dessas pessoas**. Muitas vezes, a gente está estudando grupos que são diferentes geneticamente em termos de ancestralidade; isso impacta diretamente na descoberta. A gente sabe que muitas vezes quando a gente encontra a associação de um gene com uma doença, a gente tem que **fazer um controle para ancestralidade genética** dessas populações que a gente está estudando. Porque, muitas vezes, essa diferença de frequência ela só pode ser só um reflexo da diferença étnica dessas pessoas. Nesse tipo de estudo de larga escala é muito importante que a gente faça um controle de ancestralidade; há formas metodológicas, estatísticas, de se fazer isso, para que a gente possa controlar essas associações, evitando associações espúrias.

## Populações de estudo

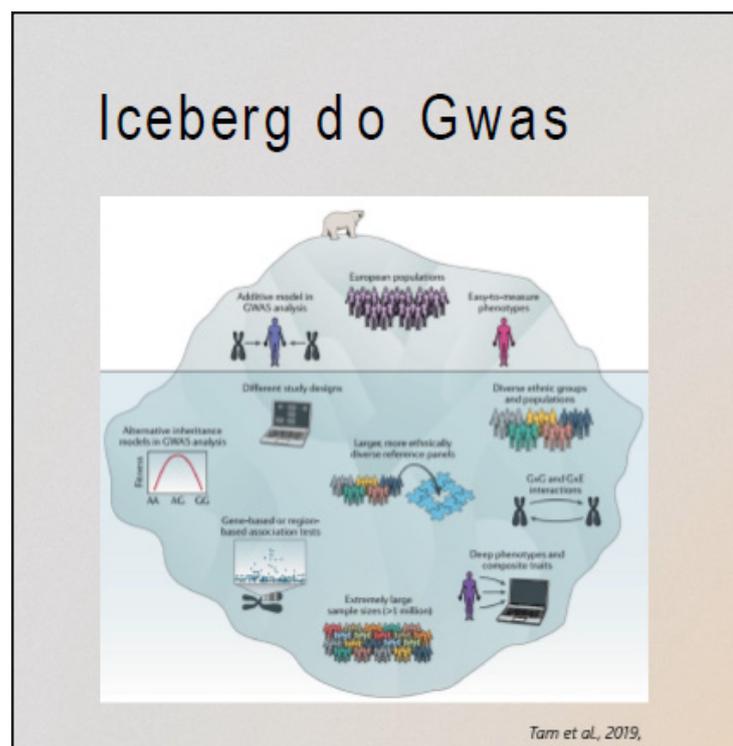
Estratificação populacional

- Grande problema
  - Grupos diferem quanto à ancestralidade genética
    - Frequência entre grupos
  - Confusão entre ancestralidade e doença
    - Uma associação espúria pode ocorrer
    - Consequência: aumento de associações falso-positivas
- Controlando para estratificação populacional é extremamente importante

Outro problema metodológico é o **tipo de etnia** encontrado nesses estudos. A gente observa nesse paper da Revista Cell que saiu em 2019, mostrando aqui todos os estudos relacionados às doenças mentais, estudos em larga escala de Gwas em larga escala para transtornos mentais - aqui Depressão, Esquizofrenia, Bipolar, Autismo - e a gente observa aqui que, da maioria dessas doenças mais de 90% das amostras eram amostras europeias. O que a gente tem na literatura de conhecimento de genes associados a essas doenças é, basicamente, nas populações europeias.



Esse outro gráfico nos mostra, mais ou menos, a mesma coisa; há uma grande desproporção de estudos com populações europeias e não europeias.



A gente precisa melhorar bastante esse tipo de estudo em larga escala, tentando controlar todas essas variáveis. Observamos esse gráfico mostrando o que a gente tem na literatura, o que está acima, o pico do iceberg, é o que está apresentado na literatura atualmente. E o que está pra baixo é o que a gente não tem ainda; a gente ainda não tem esse tipo de estudo na literatura, mas seria muito interessante que a gente fizesse, ou seja, que a gente aumentasse consideravelmente amostra; que a gente estudasse amostras que têm uma maior miscigenação ou contribuições genéticas não europeias, que a gente testasse diferentes modelos genéticos para que a gente pudesse aumentar o poder da descoberta genética desses transtornos.

## ANÁLISES PÓS-GWAS

Esses estudos em larga escala, chamados de Gwas, nos possibilitaram fazer alguns cálculos que têm nos ajudado bastante a **entender as bases genéticas desses transtornos**. Um deles é a **herdabilidade baseada em SNPs**. Na aula passada a gente viu esse conceito de herdabilidade, que é um índice que estima a contribuição genética para uma característica, para uma doença. Isso era feito com base em estudos de gêmeos. Agora a gente consegue fazer o **cálculo da herdabilidade** com base nas variantes do tipo SNPs. A gente pode estimar a herdabilidade com base em todos os SNPs testados no Gwas. A gente pode formular um índice e com isso a gente vai saber o quê?

## HERDABILIDADE BASEADA EM SNP

Qual é a proporção fenotípica, **a proporção do fenótipo que pode ser explicada apenas pelos SNPs?**



Vocês podem encontrar isso na literatura com base na herdabilidade, baseada em SNPs, que essa herdabilidade é sempre um pouco menor do que a herdabilidade - um pouco, às vezes muito aqui no caso - menor do que a herdabilidade baseada em gêmeos. Um exemplo aqui é a Esquizofrenia, a herdabilidade baseada em gêmeos da Esquizofrenia, é estimada em 80%; e a herdabilidade baseada em SNPs foi estimada para Esquizofrenia em 24%. O que isso quer dizer? Isso quer dizer que **cerca de 80% da etiologia da Esquizofrenia é atribuída a causas genéticas** e, dentro dessas causas genéticas, 24% podem ser explicados por variantes do tipo SNP. Os SNPs contribuem para 24% do fenótipo de esquizofrenia.

## Herdabilidade baseada em snp

Esquizofrenia

$h^2 = 80\%$

SNP- $h^2 = 24\%$

Esse cálculo da herdabilidade baseada em SNPs, nos permitiu estimar as correlações genéticas entre alguns transtornos mentais.

## ESTIMATIVA DAS CORRELAÇÕES GENÉTICAS

Esse *paper* da Cell, o objetivo dele era justamente esse, compreender um pouco melhor o compartilhamento genético entre os transtornos mentais com base na herdabilidade baseada em SNPs.

**Article**

**Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders**

Cross-Disorder Group of the Psychiatric Genomics Consortium<sup>1,2,3,4</sup>  
 Lead Contact: J. Smoller  
 Correspondence: [j.smoller@mgh.harvard.edu](mailto:j.smoller@mgh.harvard.edu) or [pleec@amg.harvard.edu](mailto:pleec@amg.harvard.edu)  
<https://doi.org/10.1016/j.cell.2019.03.016>

**SUMMARY**

Genetic influences on psychiatric disorders transcend diagnostic boundaries, suggesting substantial pleiotropy of contributing loci. However, the relative and mechanistic roles of these pleiotropic effects remain unclear. We performed analyses of 232,956 cases and 684,142 controls from genome-wide studies of anorexia nervosa, attention-deficit/hyperactivity disorder, autism spectrum disorder, bipolar disorder, major depression, obsessive-compulsive disorder, schizophrenia, and Tourette/tic disorder. Genetic correlation analyses revealed a meaningful structure across the eight disorders, identifying three groups of trait-related disorders. Meta-analysis across these eight disorders detected 129 loci associated with at least two psychiatric disorders, including 23 loci with pleiotropic effects on four or more disorders and 11 loci with antagonistic effects on multiple disorders. The pleiotropic loci associated with genes that show heightened expression in the brain throughout the lifespan, beginning primarily in the second trimester, and play prominent roles in neurodevelopmental processes. These findings have important implications for psychiatric nosology, drug development, and risk prediction.

**INTRODUCTION**

Psychiatric disorders affect more than 25% of the population in any given year and are a leading cause of worldwide disability (1). Since their first systematic description (2), DSM (3), ICD (4), and DSM-5 (5), the substantial evidence of genetic variation on loci for a broad range of psychiatric disorders has been established by both twin and, more recently, large-scale genomic studies (6–10). Psychiatric disorders are highly polygenic, with a large proportion of liability attributable to common variation. Many loci that have emerged from genome-wide association studies (GWAS) of among others, anorexia nervosa (AN), bipolar disorder (BIP), major depression (MDE), and schizophrenia (SCZ), have also been associated with other psychiatric disorders (11–14).

Our study used genome-wide single nucleotide polymorphism (SNP) data for eight neuropsychiatric disorders using a cross-disorder sample of 232,956 cases and 684,142 controls (Table 1). Table 1. The eight disorders included AN (Cassani et al., 2015) and DSM (Stone et al., 2018), MDE (Demirkan et al., 2018), BIP (Wray et al., 2018), SCZ (Wray et al., 2018), OCD (OCD International Consortium Disease Foundation Database Catalogue (OCDP-DC) and OCD Consortium Research Association Database (OCDRA), 2018), TS (Wray et al., 2018), and BCD (British Biobank Working Group of the Psychiatric Genomics Consortium, 2018). All study participants were of self-identified European ancestry, which was supported by principal component analysis of genome-wide data.

Cell 178 1489–1493, December 12, 2019 © 2019 Elsevier Inc. 1489

**Cell**

**Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders**

**Graphical Abstract**

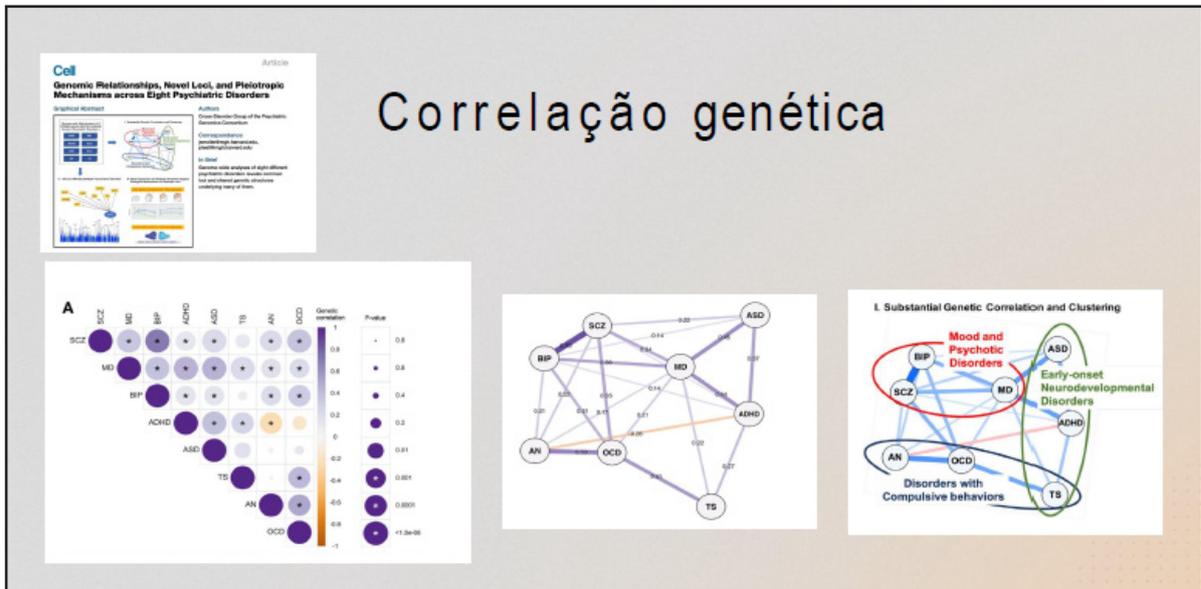
**Authors**  
 Cross-Disorder Group of the Psychiatric Genomics Consortium

**Correspondence**  
[j.smoller@mgh.harvard.edu](mailto:j.smoller@mgh.harvard.edu),  
[pleec@amg.harvard.edu](mailto:pleec@amg.harvard.edu)

**In Brief**  
 Genome-wide analyses of eight different psychiatric disorders reveals common loci and shared genetic structures underlying many of them.

A gente observa aqui nessa nesse gráfico, nessa matriz de correlação, as oito doenças estudadas neste artigo Esquizofrenia, Bipolar, Depressão, TOC, e a gente observa aqui que essa base genética que são os SNPs correlacionados ditos, foram correlacionados entre si para essas doenças e, em alguns casos, a gente encontrou um **alto compartilhamento genético entre as doenças**.

A gente observa aqui, por exemplo, entre a **Esquizofrenia e o Transtorno Bipolar**, há uma **grande correlação genética entre eles estimada em cerca de 70%**; 70% da base genética do Transtorno Bipolar é similar à base genética da Esquizofrenia. Isso aconteceu para outras doenças também; isso aconteceu para depressão e TDAH, por exemplo; ou para TDAH e Síndrome de Tourette. Com base nessa herdabilidade baseada em SNPs, a gente consegue estimar o quanto as doenças estão relacionadas geneticamente.



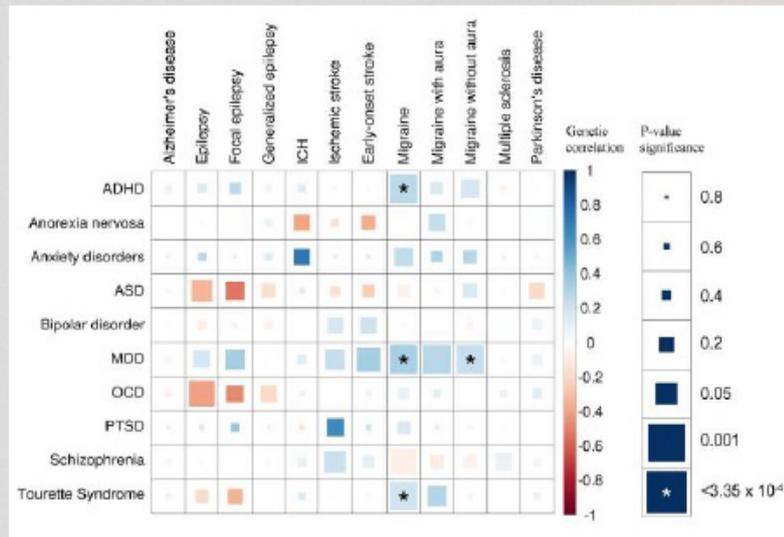
E, com isso, a gente observa que a correlação genética entre essas doenças, muitas vezes transcendem os limites diagnósticos atuais. Se a gente observa uma correlação genética de cerca de 70% entre Esquizofrenia e Bipolar, a gente observa que elas são bastante semelhantes. **Muitas vezes, a base genética é muito semelhante entre doenças que são classificadas como diferentes atualmente** pela nosologia atual dentro da Psiquiatria. Embora a Nosologia atual nos mostre essas doenças colocadas em categorias e classificações diferentes, a gente observa, dentro dos limites genéticos, que esses limites são borrados; a gente tem um grande compartilhamento genético entre as doenças. Isso acontece, por exemplo, para Esquizofrenia com várias outras doenças. **A esquizofrenia além da forte correlação com o Bipolar, ela tem uma grande correlação com Autismo, com Anorexia, com TDAH e com Depressão maior.**

## Correlação genética

Compartilhamento genético

- Influências genéticas sobre as doenças psiquiátricas transcendem limites diagnósticos atuais
- Correlação genética significativa entre esquizofrenia e 6 doenças:
  - Bipolar, DM, TDHA, TOC, TEA e anorexia
  - Forte correlação (70%) entre esquizofrenia e doença bipolar

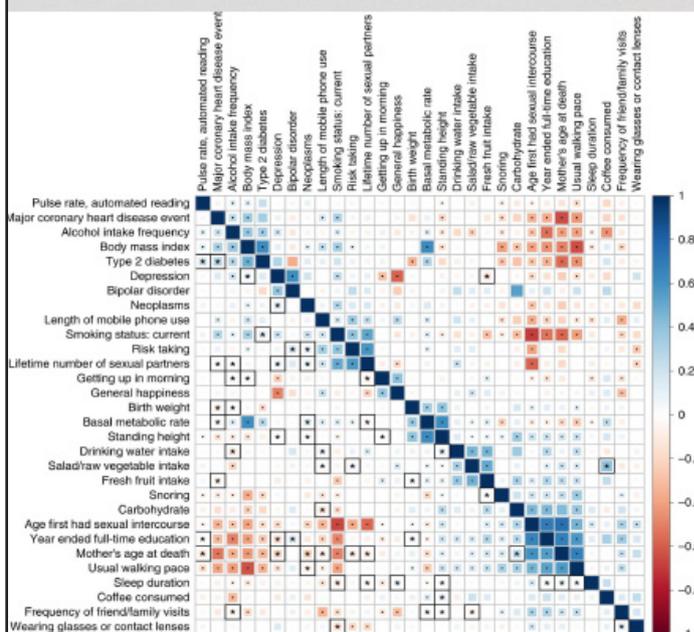
## Estimativa das correlações genéticas



Antilla et al., 2018

Alguns pesquisadores tentaram verificar se essa correlação genética era vista nos transtornos mentais, se ela também era vista nos transtornos neurológicos, ou seja, se essa correlação genética era própria de doenças do cérebro. E, para surpresa de todos, isso não aconteceu. **As doenças neurológicas não apresentam uma grande correlação genética** entre si e nem há uma grande correlação genética entre elas e os transtornos psiquiátricos, mostrando que esse compartilhamento genético é um compartilhamento próprio dos transtornos mentais, do comportamento.

## Estimativa das correlações genéticas



Ning et al., 2020, Nature genet

Outro *paper* um pouco mais recente de 2020 da Nature Genetics tentou fazer a mesma metodologia, tentando correlacionar a herdabilidade entre doenças e entre hábitos e comportamentos, para tentar correlacionar alguns fenótipos que pudessem nos ajudar a prever ou tentar compreender a etiologia de algumas doenças. Isso foi possível com base nessa matriz de correlação, correlacionando a herdabilidade das doenças e de alguns hábitos como beber, a ingestão de água, a ingestão de café ou fumar, e a correlação genética com a com algumas doenças. Certo.

## ESCORE POLIGÊNICO DE RISCO

Aqui além da herdabilidade baseada em SNPs, há outro cálculo que é possível de ser feito com base nos estudos de D Gwas, que nós chamamos hoje de Escore Poligênico de Risco.



O que é o Escore Poligênico de Risco? Vocês já podem ter ouvido falar; têm até empresas que oferecem esse tipo de serviço, que é o cálculo de Escore Poligênico de Risco. Esse escore nada mais é do que um número arbitrário, baseado na **soma de milhões de alelos de riscos genéticos** ponderados pelo seu tamanho de efeito. Eu vou explicar um pouquinho melhor o que é isso e **esse escore é individual**. Cada um apresenta o seu próprio escore de risco para Esquizofrenia ou para o Transtorno Bipolar ou até mesmo para outras doenças como Obesidade, Câncer etc.

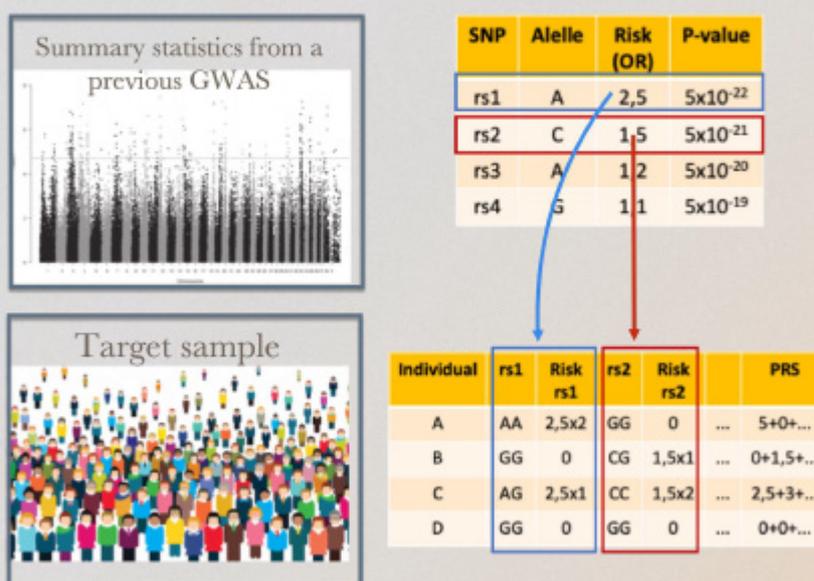
Para se calcular o Escore Poligênico de Risco de uma pessoa, você precisa ter como referência um estudo de Gwas anterior. No caso dos transtornos mentais, esses estudos que são utilizados para calcular os Escore Poligênico de Risco, na sua vasta maioria, são os Genome-Wide Association Study do Consórcio Mundial de Genômica Psiquiátrica.

## Escore poligênico de risco

- Usa como referência um GWAS: maioria é do PGC
- Escolha do melhor modelo: maior variância explicada encontrada
- Em seguida, testar na mostra-alvo

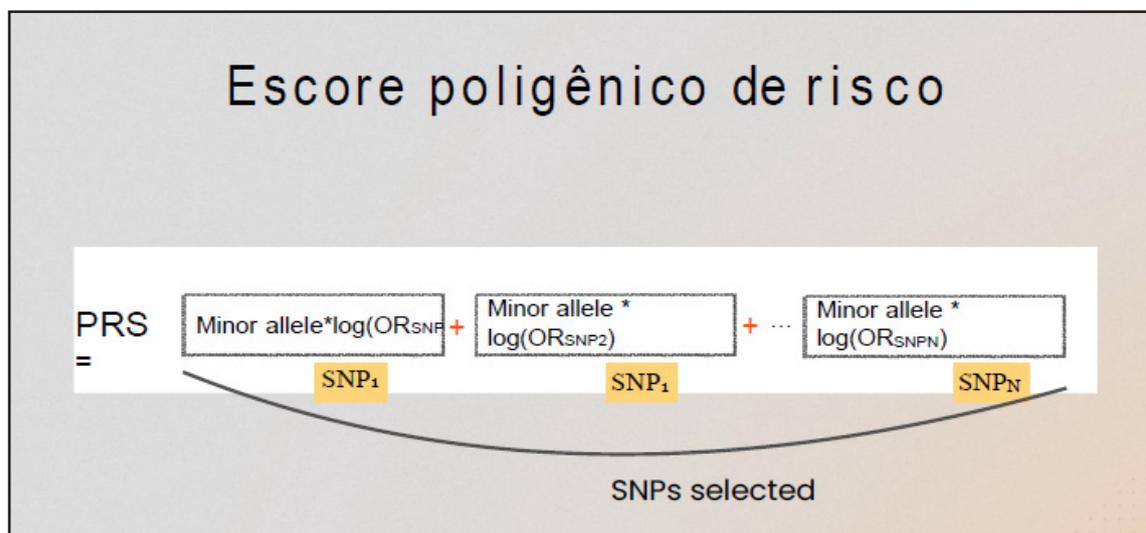
Então, nada mais é do que você pegar os resultados de um Gwas anterior, verificar qual é o tamanho de efeito para cada variante que a gente consegue através desse tipo de estudo e aplicar isso numa amostra alvo, individualmente. E aí você **testando o genótipo das pessoas, você consegue saber quantos alelos de risco ela tem para aquela doença**; pondera a quantidade de alelos de risco que ela tem para o tamanho de efeito de cada variante e, com isso, você tem uma soma, um número, que é o próprio Escore Poligênico de Risco.

## Escore poligênico de risco

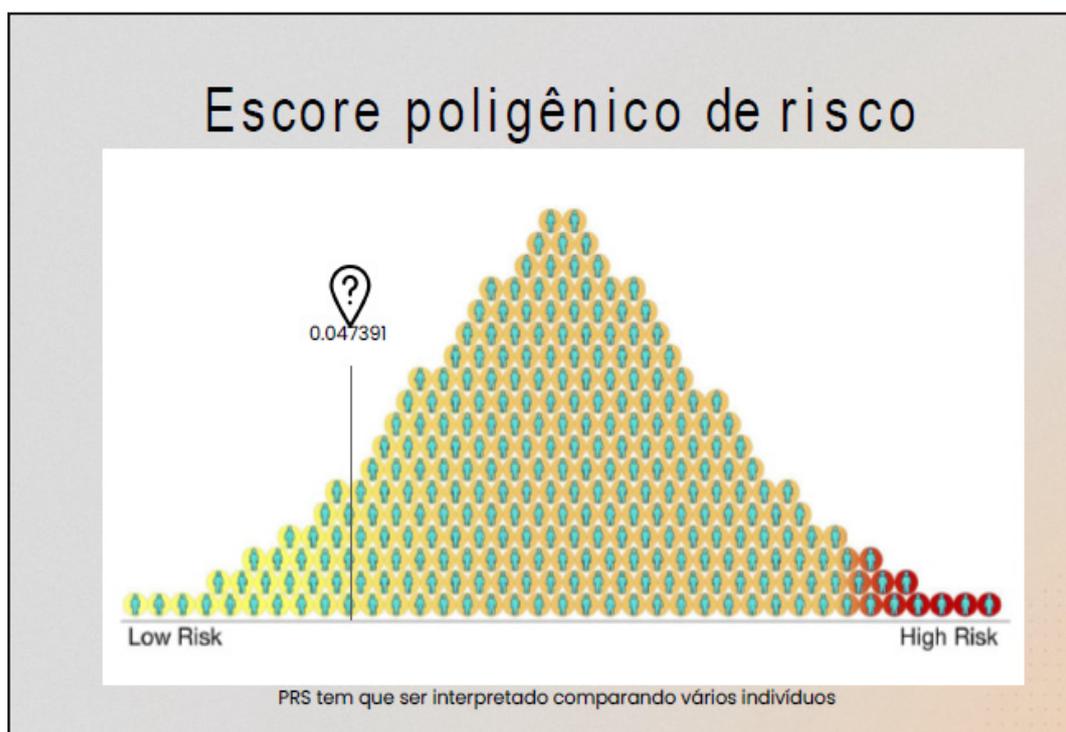


No Escore Poligênico de Risco, é preciso ter muita cautela para interpretação dele, porque é um número aleatório que **não quer dizer nada isoladamente; tem que ser interpretado no contexto de uma população**, na qual você observa se aquele número, daquele sujeito, está colocado em uma curva, em alguma das extremidades ou na média da curva.

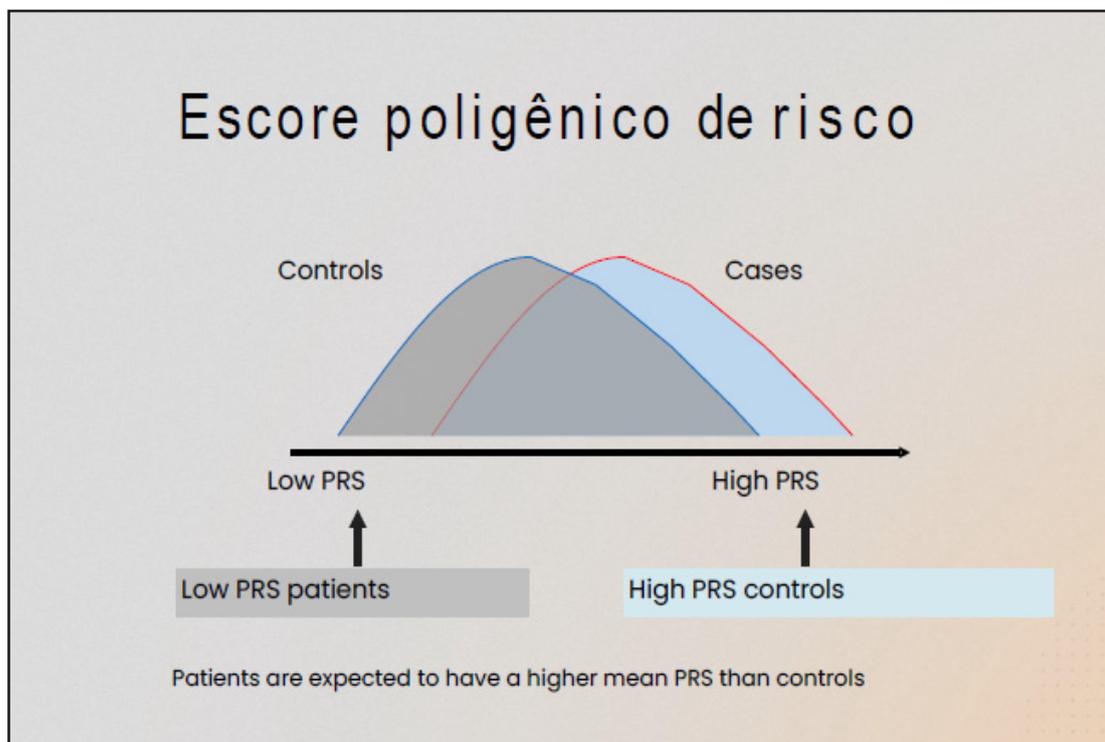
Será que esse Escore Poligênico de Risco é mais baixo do que a média da população? Será que é mais alto do que a média da população? Este *score* precisa ser interpretado no contexto populacional; não é um número que possa ser interpretado isoladamente.



A gente vai encontrar pessoas, por exemplo, que têm uma doença genética, um transtorno mental, por exemplo, e têm um Escore Poligênico de Risco. Assim como você pode encontrar pessoas normais que não tenham a doença, mas tenham um alto Escore Poligênico de Risco. Isso tem relação lá com o que a gente falou na aula passada, mostrando que a quantidade de risco que você herda dos seus genitores, **interage com a quantidade de fatores ambientais na qual o sujeito está exposto**. Por isso a gente também tem que ter bastante cautela, mesmo que a gente encontre pessoas com alto Escore Poligênico de Risco, ou até mesmo com um escore bastante baixo.



Esse escore ainda precisa ser melhorado em termos populacionais. Ele tem um alto poder preditivo em populações europeias. Por quê? Porque são os estudos de Genome-Wide Association, feitos até o momento, foram feitos em populações europeias. A gente conhece a frequência desses alelos nas populações europeias. Mas a gente não conhece nas populações miscigenadas e não europeias.



Com isso a gente tem um poder de predição um pouco menor nas populações de origem africanas ou não europeias. Essa acurácia precisa ser melhorada para populações miscigenadas como as do Brasil.

**Escore poligênico de risco**

O cálculo de PRS ainda necessita melhorar  
...especialmente em populações misigenadas e não-Europeias

Bom, com base nisso a gente pode resumir um pouco do que a gente aprendeu em termos de Genética Psiquiátrica nessas aulas.

## Escore Poligênico de Risco

Perspective | Published: 29 March 2019

### Clinical use of current polygenic risk scores may exacerbate health disparities

Alicia R. Martin, Masahiro Kanai, Yoichiro Kamatani, Yukinori Okada, Benjamin M. Neale & Mark J. Daly

Population	Approximate Median Accuracy
European	0.95
American	0.65
South Asian	0.60
East Asian	0.50
African	0.25

- Accuracy relative to European population
- We need to improve performance to cover all ancestry

## LIÇÕES APRENDIDAS

## Lições aprendidas

- Nenhum gene é necessário e nenhum é suficiente para desencadear o fenótipo
- Múltiplos genes interagem com o ambiente
- Maioria das variantes tem baixo efeito sobre a doença
- Grande amostras e com diversidade são necessárias para aumentar o poder das descobertas

O que a gente pode recapitular em termos de Genética Psiquiátrica? É importante a gente frisar que nenhum gene, **nenhuma variante é necessária ou suficiente para desencadear um único fenótipo**. Por quê? Porque múltiplos genes interagem entre si e ainda interagem com o ambiente. A maioria dessas variantes têm um baixo tamanho de efeito; isso precisa ser interpretado no contexto junto com a exposição ambiental. A gente precisa de grandes amostras e com grande diversidade genética para poder aumentar o poder da descoberta das doenças mentais.

## Lições aprendidas

- Ainda não há marcadores genéticos que dêem suporte para a atual classificação
- Há um grande compartilhamento genético entre as doenças e isso ultrapassa os limites da classificação de doenças
- Herdabilidade baseada em gêmeos é maior que a herdabilidade baseada em SNPs
- PRS pode auxiliar na predição, mas ainda só usado para pesquisa

Ainda não há marcadores genéticos que deem suporte para a classificação atual das doenças. Outra coisa importante é que **há um grande compartilhamento genético entre as doenças mentais**; esse compartilhamento ultrapassa os limites da Nosologia atual. Outro ponto que é preciso frisar é que a herdabilidade baseada em gêmeos ela é sempre maior do que a herdabilidade baseada em SNPs, porque os SNPs são uma única categoria genética dentro de todas as variantes genéticas e o Escore Poligênico de Risco é um escore que vem sendo bastante utilizado para tentar prever a doença ou as doenças complexas e pode auxiliar nessa predição, por enquanto, ele ainda é utilizado apenas em pesquisa, principalmente em populações não europeias cuja acurácia ainda é menor do que a desejada.

Era isso, pessoal, obrigada e até a próxima.