

Universidade de Évora
Departamento de Biologia

Disciplina de Bases Biológicas do Comportamento Humano

**APONTAMENTOS DA AULA DE GENÉTICA,
3 DE MARÇO DE 2005**

com correcções feitas em Novembro de 2006

Paulo de Oliveira

Fenótipos e genótipos

Fenótipos são características observáveis, por exemplo o resultado dum teste psicológico é um fenótipo. Os genótipos, que descrevem os genes que determinam os fenótipos, não são observáveis, mas podem em certos casos ser deduzidos pela análise genética. O conhecimento do genótipo dum determinado indivíduo é por sua vez imprescindível para fazer previsões sobre os fenótipos da descendência que ele pode produzir.

A análise genética clássica baseia-se nas leis de Mendel, que estipulam regras de distribuição dos genes pelos gâmetas, para determinar genótipos. Para isso, é necessária informação sobre a distribuição dos indivíduos por diferentes classes de fenótipos (cores, formas, texturas, patologias, etc.) e sobre o seu parentesco entre si. Se for possível controlar os cruzamentos (não na espécie humana, evidentemente), é mais fácil determinar o genótipo de cada indivíduo através da descendência que ele produz.

Mesmo assim, nem sempre é possível determinar os genótipos com precisão: ou porque a informação não é suficiente, ou porque os genótipos não se distribuem em classes disjuntas. Neste caso fala-se de variação contínua, ou de “caracteres quantitativos”, donde se pode quando muito estimar até que ponto as diferenças entre indivíduos se devem aos seus genótipos ou a outras causas. As características do comportamento humano têm de ser todas abordadas por esta via, mesmo no caso das psicopatologias, apesar de nestas haver uma separação objectiva de duas classes (presença vs. ausência da patologia) — só que essa separação é fenotípica apenas, não se lhe consegue atribuir uma formulação genotípica.

Análise da variação contínua

A teoria genética da variação contínua parte do princípio que a distribuição dos fenótipos se aproxima da gaussiana (normal), o que se verifica em muitos casos, incluindo os fenótipos comportamentais avaliados por testes psicológicos e inquéritos. Os parâmetros que descrevem esta distribuição são a média, o desvio-padrão, e o número de observações.

O resultado dum indivíduo num teste psicológico é um valor na escala de variação admitida por esse teste, que designaremos por P (inicial de *phenotype*). A questão fundamental na Genética da variação contínua é determinar a contribuição do valor genotípico G (de

genotype), e das influências ambientais E (de *environment*), sobre o desenvolvimento do fenótipo P.

$$P = G + E$$

G é essencialmente o efeito acumulado, ou concorrente, de um grande número de genes, geralmente distribuídos um pouco por todos os cromossomas, dando cada um uma contribuição modesta; diz-se por isso que as características de variação contínua são poligénicas. O valor de E é a resultante de influências não-genóticas, positivas ou negativas, que modificam o efeito dos genes, ou seja, podem desviar P de G. Pela possibilidade de substituição mútua dos componentes individuais tanto de G como de E, decorre a noção que dois indivíduos podem ter o mesmo P com valores G muito diferentes entre si; por isso, G e E são duas incógnitas irresolúveis.

A análise genética tem de recorrer à distribuição no seu todo e centrar-se no estudo das causas de variação de P entre os indivíduos no seu conjunto. O parâmetro da distribuição gaussiana relevante para esse estudo é o quadrado do desvio-padrão de P, que se designa por variância fenotípica, V_P . Este parâmetro decompõe-se numa variância de valores genotípicos V_G , numa variância de influências ambientais V_E , e numa covariância genótipo-ambiente COV_{GE} :

$$V_P = V_G + V_E + 2COV_{GE}$$

Assim, V_G é uma medida das diferenças genóticas entre os indivíduos, ou seja, é a variância duma distribuição, subjacente, dos valores G; a V_E é-lhe análoga, mas em referência às influências ambientais E; quanto à COV_{GE} , é uma medida da interdependência entre os valores genotípicos e a respectiva susceptibilidade às influências ambientais, pois existem características em que essa susceptibilidade pode ser maior ou menor consoante o genótipo de cada indivíduo.

Assim, quando se coloca a questão de saber qual a componente do fenótipo que é genotípica, a resposta não se faz directamente, mas sim através da proporção das diferenças fenotípicas, representadas pela V_P , que é devida a diferenças entre genótipos, portanto, do quociente V_G/V_P . A restante proporção é atribuída a outras causas, genericamente designadas influências ambientais (embora também se deva levar em conta a sua interdependência com os genótipos, representada pela componente COV_{GE}).

Mas pode ir-se além da separação dum componente genotípica da variação fenotípica. Desde que haja informação suficiente, podem separar-se componentes tanto de G como de E. A teoria genética postula que existe uma componente aditiva, A, outra de dominância, D, e outra de epistasia (ou, mais rigorosamente, de interacção entre genes não-alelos), I, tal que

$$V_G = V_A + V_D + V_I$$

A componente V_A é a que se identifica como a proporção de V_P que determina as diferenças entre descendências dos diferentes indivíduos, o que é especialmente importante, por exemplo, para o melhoramento de espécies domésticas por selecção artificial (trata-se aqui de prever até que ponto se pode obter uma descendência “melhor” a partir da selecção de progenitores fenotipicamente “melhores”). Daí que a V_A seja a variância da distribuição, subjacente, de valores melhoradores A. Por definição, a dependência dos fenótipos de cada descendência D em relação a um dos progenitores, medida através do coeficiente de regressão (b_{DP} , semelhante ao de correlação, mas não idêntico) é uma estimativa de V_A ¹. Designa-se colectivamente a soma $V_D + V_I$ como variância genotípica não-aditiva (V_{NA}), e nela a parcela V_D é em geral a mais relevante.

A proporção V_A/V_P designa-se heritabilidade (aportuguesamento do inglês *heritability*), simbolizada por h^2 . Este parâmetro tem uma grande importância nos estudos sobre evolução biológica, assim como no contexto aplicado do melhoramento de plantas ou animais, mas não é especialmente relevante para a análise genética do comportamento humano; é a proporção V_G/V_P (às vezes designada heritabilidade em sentido lato, embora discutivelmente) que tem a maior relevância para este campo, pois através dela responde-se à questão essencial sobre a contribuição relativa das diferenças entre genótipos para a variância fenotípica.

As componentes de E são em teoria inúmeras, tudo dependendo da quantidade de informação disponível sobre as diversas influências ambientais; assume especial relevância a identificação de causas ambientais partilhadas, quer sejam elas pré-natais (no útero, e aqui pode até discriminar-se, no caso de gémeos monozigóticos, entre os monoplacentários e diplacentários) ou pós-natais, designadamente o ambiente familiar (nutrição, protecção de doenças, comportamento dos pais, educação...). Se considerarmos as pós-natais como uma

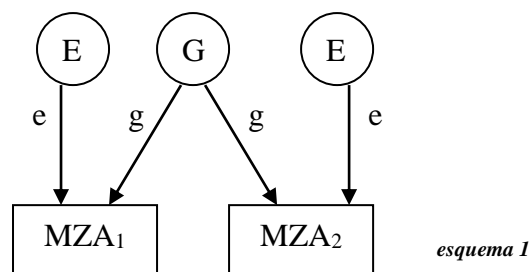
¹ $V_A = 2b_{DP}V_P$. O significado de P nesta equação é, em ambos os factores, o da distribuição fenotípica da população de progenitores. Dentro de certos limites e precauções, esta medida de interrelação entre gerações pode adequar-se ao estudo de características físicas, porém já o mesmo não se pode dizer das características de comportamento humano.

contribuição C para o fenótipo P, separa-se então uma componente V_C da variância fenotípica, referente às diferenças entre influências ambientais que distinguem os agregados familiares, sendo o restante ainda designado como V_E (componente de causas não identificadas da V_P), que neste contexto significa variância de influências ambientais não-partilhadas.

A componente COV_{GE} não é em geral quantificada, ficando englobada pela V_E .

Metodologia

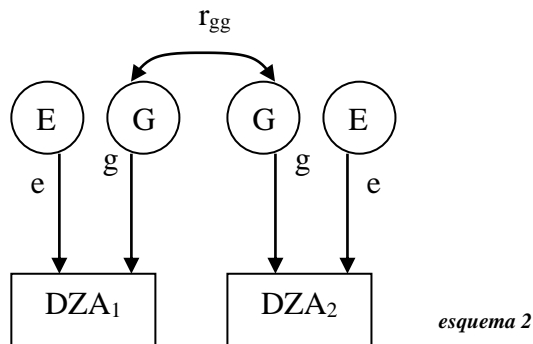
Só através de correlações dentro da componente genotípica e/ou da componente ambiental se podem fazer estimativas fiáveis das componentes da variância fenotípica. Na espécie humana, a abordagem mais útil baseia-se nos estudos em gémeos e adopções. Os membros dum par de gémeos monozigóticos (MZ) têm genótipos virtualmente idênticos, por isso qualquer diferença entre eles tem de ser atribuída a outras causas; se foram dizigóticos (DZ), têm uma correlação genotípica teoricamente igual à dum outro par de irmãos mas partilhando pelo menos em parte o ambiente uterino. Os diagramas de causalidade ilustram convenientemente as componentes em jogo dentro de cada modelo de parentesco. Examinando o caso de dois gémeos monozigóticos criados separadamente (MZA), esquema 1:



Os rectângulos representam os fenótipos (por exemplo, os desempenhos num mesmo teste psicológico) dos gémeos MZA_1 e MZA_2 ; os círculos representam os valores ambientais (E) e o valor genotípico (um único G) que determinaram esses fenótipos; finalmente, as setas indicam correlações entre causas e efeitos, isto é, o grau de influência dum factor causal (G ou E) sobre a manifestação genotípica, e designam-se g ou e, consoante a ligação é com a causa genotípica ou a ambiental, respectivamente². Nesta situação, o investigador apenas consegue identificar um mesmo valor g para ambos os gémeos, donde a correlação entre fenótipos de pares de gémeos r_{MZA} , medida para o conjunto da população, é $r_{MZA} = g \times g = g^2$.

² Blouchard & McGue (2003) utilizam a letra h, para identificá-la com a heritabilidade, mas a componente do genótipo que é expressa no fenótipo não é o seu valor melhorador A, mas sim o seu valor genotípico G, donde resulta que a variância que está sempre em causa é a V_G e não a V_A .

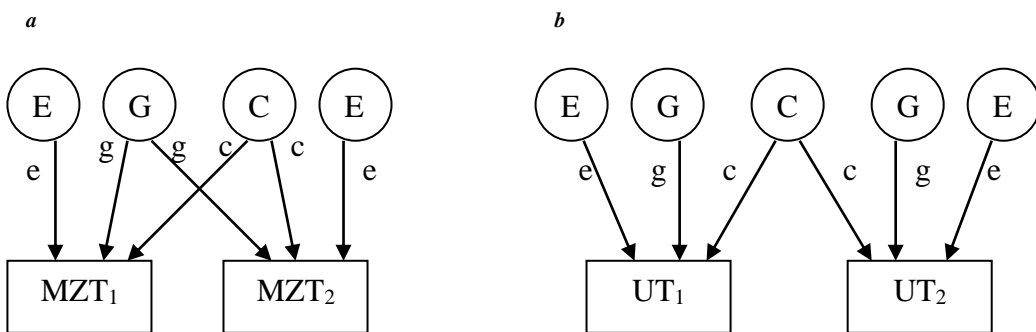
É importante notar que, sendo g^2 uma medida aparentemente directa da componente genotípica da variância fenotípica (isto é, uma estimativa de V_G/V_P), na realidade há que produzir uma medida independente de g^2 que, em condições análogas, dê uma correlação diferente. Isso deve-se principalmente ao facto de haver muitos factores em comum entre gémeos (desenvolvimento no mesmo útero, nomeadamente) que aumenta a correlação entre eles, mesmo que depois sejam criados em separado. Os gémeos DZA (criados separadamente) fornecem o necessário ponto de referência (esquema 2):



Neste caso, $r_{DZA} = g^2 r_{gg}$, sendo r_{gg} a correlação genotípica entre irmãos que, ignorando V_I , é igual a $(\frac{1}{2} V_A + \frac{1}{4} V_D)/V_G$. Assim, obtém-se $g^2 = (r_{MZA} - r_{DZA})/(1 - r_{gg})$, e se se assumir $V_D = 0$, $V_G = V_A$, donde $r_{gg} = \frac{1}{2}$ e $g^2 = 2(r_{MZA} - r_{DZA})$, fórmula esta muito utilizada como aproximação de V_G/V_P .

Uma estratégia semelhante pode desenhar-se na comparação entre gémeos monozigóticos criados juntos (MZT) e entre indivíduos não-aparentados criados juntos (UT, por exemplo um filho biológico e outro adoptado), que permite o cálculo de c^2 , derivado da componente ambiental C (esquema 3).

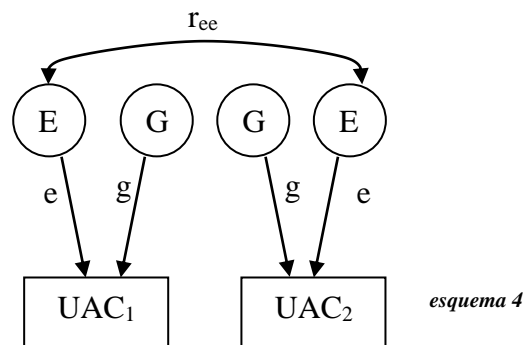
esquema 3



Assumindo que os valores G e C são independentes, a correlação r_{MZT} entre os fenótipos dos pares de gémeos é $g^2 + c^2$, e r_{UT} é c^2 .

Note-se que a diferença $r_{MZT} - r_{MZA}$ corresponde a c^2 , permitindo assim avaliar-se se a fórmula aproximativa de g^2 , assumindo $V_D = 0$ como explicado acima, é aplicável.

Uma variante do esquema 3b é a de indivíduos não-aparentados criados separadamente mas em ambientes correlacionados (UAC), por exemplo em situações socialmente características (bairros degradados, colégios internos dirigidos por religiosos, etc.). Aqui, o que deveria ser uma correlação nula pode ser positiva, em função dessa correlação ambiental, expressa pelo parâmetro r_{ee} (esquema 4):



Analogamente aos exemplos anteriores, $r_{UAC} = e^2 r_{ee}$.

Problemas com o desenho da investigação genética

Nem sempre as correlações são fáceis de interpretar. Sem haver uma relação clara de causa-efeito entre acções genóticas e/ou ambientais e os valores fenotípicos, pode estar-se a deduzir uma relação que não tem significado. A complexidade das influências familiares sobre o desenvolvimento dos caracteres comportamentais torna qualquer estudo demasiado susceptível a críticas de forma, que só se previnem com uma criteriosa amostragem de modo a assegurar algum controlo sobre as causas de variação.

Um exemplo de falha bastante comum é a tendência geralmente involuntária de excluir casos demasiado extremos, como por exemplo valores fenotípicos muito altos ou muito baixos, discrepâncias excessivas entre dois gémeos (dum mesmo par, entenda-se), etc.. Chama-se a isto truncar a distribuição fenotípica, e embora pareça não afectar os valores médios, pode ter grandes implicações sobre a variância fenotípica (e sobre as correlações). Por muito “patológicos” que tais casos pareçam, se não existirem informações adicionais que permitam considerá-los como anómalos (por exemplo, psicopatologias) não é legítimo excluí-los da análise.

Outro exemplo de falha é o enviezamento de certos factores causais. Se as famílias adoptivas pertencerem sobretudo a um estrato social elevado, com grande controlo sobre as condições de desenvolvimento da criança adoptada, a amostragem não é representativa das condições ambientais (para filhos biológicos, assim como adoptivos) na população em geral.

Interpretação da causalidade genotípica dos caracteres comportamentais

O facto de haver uma componente genotípica na variação comportamental tem várias consequências. Pode servir para considerações de carácter eugénico (utopia de “melhoramento” da espécie humana, pela limitação da reprodução de indivíduos com fenótipos considerados indesejáveis), que conduziram a excessos em regimes como o da Alemanha nazi, e na sequência disso, por reacção, ao extremo oposto de não admitir-se essa componente; mas também pode servir para compreender a Humanidade através duma vasta diversidade genética para as características comportamentais, cujo reflexo é a múltipla complementaridade entre as aptidões e inclinações que se observam a todos os níveis de integração social. Se encarada como uma mais-valia adaptativa concomitante ao processo evolutivo que desembocou na espécie humana actual, esta diversidade ganha uma dimensão histórica e biológica que a Psicologia não deve menosprezar. Além do mais, a perspectiva evolutiva permite radicar os padrões comportamentais que se observam na nossa espécie num referencial mais vasto de parentescos com outras espécies de Primatas ou outros mamíferos que, como modelos de observação ou experimentais, podem trazer importantes contributos.

A observação que o gémeos MZ têm maior tendência a partilhar amizades e a receberem o mesmo tratamento parental, quando comparados com os DZ, é um exemplo entre muitos dos problemas de interpretação que dividem os psicólogos. Para uns, esta observação invalida o postulado que a correlação de valores ambientais partilhados (componente C do fenótipo) seja igual para os dois tipos de gémeos; mas para outros, apenas indica que as diferenças genotípicas nos DZ são suficientes para a expressão de diferenças comportamentais (escolha dos amigos) e seus reflexos no comportamento dos outros (tratamento parental).

Genética dos domínios normais do comportamento humano

Capacidades cognitivas

As capacidades cognitivas são classificadas numa série de factores de primeira ordem³, para as quais existem testes específicos que permitem quantificá-las e, por isso, estabelecer o fenótipo de cada indivíduo. Acontece que estes factores de 1ª ordem não são independentes entre si, como o indica a correlação entre os seus valores dentro do mesmo indivíduo. Esta constatação leva a considerar-se que o factor principal de variação nas capacidades cognitivas se encontra a um nível hierárquico superior aos factores de 1ª ordem, um “factor g” (inteligência geral) cuja medição é o objecto dos testes de QI. Por outras palavras, a influência recíproca entre os diferentes factores de 1ª ordem é mediada pelo factor g, que lhes serve de interligação. Este esquema hierárquico é considerado válido, independentemente da etnicidade, para todas as sociedades industrializadas, e admite-se que a variação do factor g se correlaciona com certas características anatómicas do cérebro.

Os estudos com gémeos e de adopção revelam uma notável consistência nas estimativas da V_G para as capacidades cognitivas, aproximadamente 50% da V_P . As discrepâncias observadas entre o peso relativo das componentes não-genotípicas pode dever-se à idade em que é feita a observação, pois a V_C tende para 0 a partir dos 12 anos, como é evidenciado por estudos longitudinais (isto é, observando os mesmos indivíduos ao longo do seu desenvolvimento pós-natal, até à idade adulta). A própria proporção de V_G em relação à V_P também varia com a idade, só estabilizando no valor consensual de 50% depois dos 4-5 anos — presume-se que antes dessa fase a expressão da componente genotípica das capacidades cognitivas é incipiente.

Personalidade

Apesar de não haver consenso quanto à medição deste domínio do comportamento, pelas discrepâncias entre diversas teorias, a V_G tende a situar-se nos 40 a 55% da V_P , e a V_C é, de maneira consistente, nula.

³ Capacidades mentais especiais: fluidez de raciocínio, abrangência de conhecimentos, processamento visual e auditivo, velocidade de processamento, memória de curto prazo, recordações de longo prazo, capacidades quantitativas

Atitudes sociais

Resultados semelhantes aos dos caracteres de personalidade para a religiosidade, o conservadorismo e o autoritarismo. Ao mesmo tempo, a filiação religiosa é quase inteiramente determinada pela componente ambiental não-partilhada (E).

Interesses psicológicos

Comportamentos realístico, investigativo, artístico, social, empreendedor, convencional. Grande consistência nas componentes não-genóticas (V_C aproximadamente 10% da V_P , V_E aproximadamente 50%), e uma tendência para V_D maior do que V_A (sobredominância).

Psicopatologia

Este último domínio do comportamento requer abordagens diferentes dos restantes, pois como se trata duma variação entre o que é normal e não é (esquizofrenia, autismo, alcoolismo, depressão, criminalidade, etc...), os fenótipos são disjuntos, e a variação não é aparentemente quantitativa. No entanto, os estudos familiares demonstram que a sua ocorrência é esporádica, consistente com um quadro que em Genética médica é designado como multifactorial (isto é, poligénico). O risco (*liability*) dela desenvolver-se é que tem variação contínua, com uma contribuição genotípica e outra ambiental, mas o facto da manifestação fenotípica ser descontínua não permite o cálculo de variâncias.

Continua aqui a ser fundamental estudar gémeos e, até um certo ponto, adopções. A estatística que se obtém é uma percentagem de concordância (percentagem de pares de indivíduos em que ambos desenvolvem a patologia) que, se for maior em gémeos MZ do que nos DZ, indica seguramente uma componente genotípica na variação contínua do risco. Esta situação tem sido verificada em quase todas as psicopatologias. Quanto às estimativas da componente V_C , ela tende a ser irrisória excepto nos comportamentos anti-sociais (e mesmo assim, como observado nas capacidades cognitivas, reduzindo-se com a idade).

Apesar da sua base poligénica, nestas características há uma intensa investigação no sentido de identificar pelo menos parte dos genes que determinam maior risco de desenvolvimento da patologia. A exemplo doutras áreas da genética da variação contínua, há duas abordagens complementares: a delimitação de regiões cromossómicas contendo esses genes (mapeamento de QTL, abreviatura de *quantitative trait loci*), e o estudo de genes candidatos, seleccionados pela presumível relação causa-efeito das variações alélicas nesses genes com os riscos de

patologia. Ambas têm sido muito facilitadas pelo enorme avanço que constituiu a sequenciação do genoma humano, do ratinho de laboratório, e outros modelos relevantes. Pela primeira foi possível determinar, por exemplo, que o braço curto do cromossoma 8 (8p), e os braços longos dos cromossomas 13 e 22 (13q, 22q) contêm QTL para esquizofrenia e também para a doença bipolar; genes candidatos que se têm estudado incluem os de diversos neurotransmissores (COMT para esquizofrenia, serotonina para ansiedade/ impulsividade) ou transportadores de membrana (DAT1, da dopamina, para o déficite de atenção/ hiperactividade), mas nem sempre há que procurar entre os que se actuam no sistema nervoso: certas formas mais activas do enzima álcool desidrogenase, que processa o etanol no fígado, são protectoras do desenvolvimento do alcoolismo.

A importância de E na manifestação duma patologia é exemplificada por uma variante menos activa da monoamina oxidase A, que estava associada a comportamento anti-social só em rapazes que eram maltratados, não nos que não o eram. Este enzima está envolvido no metabolismo de neurotransmissores como a serotonina, a dopamina e a norepinefrina, é o alvo molecular de vários fármacos antidepressivos, e vários alelos do respectivo gene têm sido identificados com psicopatologias de pânico, doença de Alzheimer ou agressividade. Aliás, também pode incluir-se este exemplo no conceito de interacção genótipo-ambiente (determinante da COV_{GE}).

Referências

- Bouchard, Jr., T. J., McGue, M., 2003. Genetic and Environmental Influences on Human Psychological Differences. *J Neurobiol* 54: 4-45. DOI: 10.1002/neu.10160 (cf. <http://home.uevora.pt/~oliveira/Bio/Geral.pdf>)
- Pinker, S., 2004. Why nature & nurture won't go away. *Daedalus* 133(4): 3-17. DOI: 10.5007/737 (cf. <http://home.uevora.pt/~oliveira/Bio/Pinker.pdf>)

Apêndice: principais questões epidemiológicas sobre as componentes da variância populacional de características psicológicas

- A. Componentes ambientais da variância
1. Até que ponto a característica é influenciada por factores ambientais?
 2. Que tipo de acção ambiental é envolvida?
 - a. Se pré-natal, é nutricional ou hormonal?
 - b. Se pós-natal, é física ou social?
 - c. É partilhada ou idiossincrática?
 - d. Há efeitos maternos e/ou paternos?
 - e. Há efeitos da interacção entre irmãos?
 3. Há efeitos de género?
 4. A transmissão é horizontal (por exemplo, entre parceiros de grupo) e/ou vertical (por exemplo, de pais para filhos)?
- B. Componentes genéticas da variância
1. Até que ponto a característica é influenciada por factores genéticos?
 2. Que tipo de acção genética é envolvida?
 - a. Aditiva?
 - b. Dominante?
 - c. Epistática?
 3. Quantos *loci* estão envolvidos?
 4. Há limitação a um dos sexos ou ligação a um dos cromossomas sexuais?
 5. Estão envolvidos efeitos cromossómicos?
- C. Até que ponto a variação se deve ao acaso, ou a processos caóticos que se assemelham ao acaso?
- D. Influências conjuntas, genéticas e ambientais
1. Há de todo interacções genéticas × ambientais?
 2. Há correlações entre genes e ambiente?
 - a. Passivas?
 - b. Evocativas?
 - c. Activas?
- E. Influências do desenvolvimento, incluindo o envelhecimento
1. Há diferentes genes a intervir durante o desenvolvimento?
 2. Há diferentes influências ambientais a intervir durante o desenvolvimento?
 3. As variâncias devidas às diferentes categorias de influência mudam com o tempo?
- F. Semelhança entre parceiros sexuais
1. Se presente, a selecção entre parceiros sexuais é devida à semelhança fenotípica ou a homogamia social?
 2. Há diferenças entre sexos na preferência do parceiro em função da característica?
- G. Selecção
1. Que tipos de factores selectivos estavam em jogo durante a evolução da característica?
 2. Há factores selectivos em actuação actualmente?
 3. A característica é uma adaptação?