

## Fenótipos dominantes e recessivos: distorções conceituais de professores de biologia em formação

Beatriz Ceschim e Ana Maria de Andrade Caldeira

Universidade Estadual Santa Cruz, Departamento de Ciências Biológicas, Brasil, [bceschim@uesc.br](mailto:bceschim@uesc.br); Universidade Estadual "Júlio de Mesquita Filho", Departamento de Educação, Brasil, [anacaldeira@fc.unesp.br](mailto:anacaldeira@fc.unesp.br)

**Resumo:** Conceitos de genética são distribuídos em duas disciplinas independentes (clássica e molecular) no curso de Ciências Biológicas. Sabemos que o entendimento do fenômeno da dominância completa requer compreensão conceitual que relacione os processos de transmissão do material genético (hereditariedade) e os processos de produção do fenótipo. Partindo da hipótese de que possa existir uma deficiência formativa referente à compreensão da dominância, objetivamos identificar explicações que professores em formação atribuem para o fenômeno. O instrumento de coleta foi um questionário escrito. Oriundos de dois campi de uma universidade estadual brasileira, os graduandos (n=101) já haviam cursado as disciplinas de genética. Os dados foram analisados por meio da Análise de Similitude pelo *software* estatístico IRaMuTeQ, que é sustentado pelo *software* R (versão 4.0.3); e por meio da Análise de Conteúdo (interpretação categorial). A Análise de Similitude indicou os termos e relações mais frequentes, permitindo o apontamento de uma estrutura explicativa genérica, que representa o conjunto dos dados de modo amplo. Ao contrário, a Análise de Conteúdo possibilitou a verificação das especificidades das explicações por meio da construção de categorias *a posteriori*. Graduandos não apontam corretamente mecanismos subjacentes à dominância e foi possível registrar uma pluralidade de equívocos. Implicações didáticas para o ensino de biologia referentes aos equívocos são apontadas.

**Palavras-chave:** ensino de biologia, formação de professores, dominância, genética.

**Title:** Dominant and recessive phenotypes: Physiological mechanisms present in the explanation of preservice biology teachers

**Abstract:** Genetic concepts are distributed by two independent disciplines (classical and molecular) in the Biological Sciences course. We know that understanding the phenomenon of complete dominance requires a conceptual understanding that brings together the processes of transmission of genetic material (inheritance) and the production processes of the phenotype. Based on the hypothesis that there is a formative deficiency regarding the understanding of dominance, the objective is to identify explanations that preservice teachers attribute to the phenomenon. Data were obtained from a written questionnaire with open questions. The undergraduates (n = 101) are from two campuses at a public Brazilian university and have already studied the genetics subjects. The data were analyzed using the Similitude Analysis by the statistical software IRaMuTeQ,

which is supported by the R software (version 4.0.3); and using Content Analysis (interpretation from categories). The Similitude Analysis indicated the most frequent terms and relationships, allowing the identification of a generic explanatory structure, which represents the data in a broad way. Conversely, Content Analysis allowed the verification of the particularities of the explanations based on the construction of categories *a posteriori*. The undergraduates did not correctly explain mechanisms that cause dominance and a variety of misconceptions were noted. Didactic implications are discussed.

**Keywords:** biology teaching, preservice teacher, dominance, genetics.

## **1. Introdução**

A distribuição das disciplinas de Ciências Biológicas ocorre por meio de subdivisões disciplinares, resultando na fragmentação de conteúdos, ainda que sejam intimamente relacionados, como a divisão que é feita entre a genética de transmissão e a genética molecular, por exemplo.

A genética de transmissão (ou genética clássica) é aquela que se ocupa dos estudos de hereditariedade, ou seja, o foco é o organismo individual: como é herdada a constituição genética e como é transmitido o material genético para a geração subsequente (Pierce, 2011). Já a genética molecular trata de explicações relacionadas à expressão gênica e à produção de polipeptídios. Assim, aborda-se em uma disciplina como o material genético é herdado e como se dispõe na transmissão genealógica, e, em outra aborda-se como esse material genético origina fenótipos.

Partimos da hipótese de que tal organização curricular subdivida poderia dificultar que graduandos apontassem como uma disciplina se conecta com a outra. Na desassociação das disciplinas de genética de transmissão e genética molecular encontra-se o cerne do problema da presente pesquisa: se o conteúdo de genética de transmissão é ensinado em uma disciplina e o conteúdo de genética molecular é ensinado em outra disciplina, como os estudantes conseguiriam atribuir uma explicação integrada para o fenômeno da dominância?

Se a dominância é um fenômeno estudado somente sob a perspectiva da genética de transmissão e os mecanismos moleculares de produção de fenótipos são ensinados de modo genérico (sem exemplificar com casos específicos de dominância), poderia haver uma lacuna na formação conceitual dos graduandos de biologia. Com base nessa problemática, foi delineada a pergunta de pesquisa: Como professores em formação explicam o fenômeno da dominância?

Sendo assim, foi objetivo dessa pesquisa identificar quais são as explicações fisiológicas que professores de biologia em formação atribuem para o processo de dominância completa. Com base nos resultados dessa pesquisa, esperamos contribuir com dados referentes às compreensões do fenômeno da dominância, uma vez que se pretende identificar os elementos constitutivos do discurso utilizado pelos participantes. Também se espera que seja possível inferir como a linguagem influencia o pensamento e, por consequência, a aprendizagem de genética referente à dominância.

## 2. Fundamentação teórica

### 2.1 O fenômeno da dominância completa

Gregor Mendel, observando e interferindo na reprodução de vegetais, constatou um fenômeno: ao cruzar plantas de ervilhas amarelas com plantas de ervilhas verdes, obteve somente descendentes com ervilhas amarelas. A cor verde tinha "sumido". Eis o fenômeno da dominância, o qual é caracterizado pelo desaparecimento de um dos fenótipos parentais.

Na terminologia mendeliana, a característica que está presente nos descendentes é a característica "dominante", enquanto a característica que desaparece é a "recessiva" (Reece et al., 2015).

O ensino da genética mendeliana necessariamente envolverá o uso de tais conceitos e a citação do exemplo das plantas de ervilhas. Porém, não necessariamente explicitará quais são os mecanismos subjacentes ao fenômeno do referido desaparecimento da característica recessiva.

A "genética" pode ser entendida como "o estudo científico da hereditariedade e da variação hereditária" (Reece, 2015, p. 1363) e se constitui como uma disciplina científica que teve desdobramentos em subáreas, como a "genética molecular". Durante o século XX, conceitos genéticos foram estruturados, como a distinção entre genótipo e fenótipo por Wilhelm Johannsen (Watson et al., 2015), que hoje pode ser entendida como genoma (material hereditário) *versus* a "expressão" desse genoma, originando estruturas (Nachtomy, Shavit e Yakhini, 2007).

Para nomear conceitos originados, algumas terminologias são comuns entre geneticistas, por exemplo: uma unidade que mais comumente ocorre na natureza é denominada "selvagem" (Griffiths et al., 2011). Quando se trata do genótipo, utiliza-se a mesma denominação. Assim, quando um alelo (forma alternativa de um gene) ocorre com mais frequência em uma população, o que arbitrariamente chamamos de "normal", é tomado como o alelo de tipo selvagem (Klug et al., 2010). Um alelo selvagem distingue-se de um alelo mutante, pois este último será constituído por sequências de nucleotídeos alteradas e, conseqüentemente, poderá codificar produtos gênicos alterados (Klug et al., 2010).

Em indivíduos diploides, combinações de cromossomos homólogos formam pares que podem conter dois alelos selvagens ou dois mutantes (condição homozigótica); ou então formam o par com um alelo diferente do outro em cada um dos cromossomos homólogos (condição heterozigótica). Em condição heterozigótica, diversos fenótipos podem emergir.

É no indivíduo heterozigoto que o fenômeno da dominância acontece, pois os possíveis produtos dos alelos selvagem e mutante irão interagir. Em alguns casos, essa interação pode resultar na produção de um fenótipo que seja semelhante somente a um dos indivíduos parentais, o que caracteriza a dominância completa. Como propôs Gregor Mendel:

O caráter híbrido se assemelha ao de uma das formas parentais tão intimamente que a outra escapa completamente da observação ou não pode ser detectada com clareza. [...]. Esses caracteres do híbrido serão denominados "dominantes" e aqueles que se tornam latentes no processo serão denominados "recessivos". A expressão "recessivo" foi

escolhida porque os caracteres assim designados se retiram ou desaparecem completamente nos híbridos, mas reaparecem inalterados em sua descendência [...] (Mendel, 2008 [1909], p. 15, tradução nossa).

Com base no trecho acima, é possível afirmar que as palavras metafóricas “dominante” e “recessivo” são propostas por Mendel tendo como critério um caráter fenomenológico, e não fisiológico. O monge faz uso de tais termos como adjetivos, nunca como substantivos ou utilizando um verbo correspondente (Allchin, 2000). Deste modo, nunca descreveu uma relação de dominação entre dois personagens, apenas os rotulou com base em sua aparência visível (na prole híbrida) (Allchin, 2000).

Mendel fez descrições quantitativas acerca das diferentes características que as plantas exibiam a cada geração, embora não tenha apontado quais mecanismos causavam tais diferenças. Ou seja, a genética de transmissão é tratada de modo independente da genética molecular desde a sua origem, o que se manteve inalterado também no ensino e na aprendizagem até os dias atuais.

Embora relacionadas, as duas disciplinas não têm fronteiras bem definidas para determinar onde uma começa e a outra termina. As duas disciplinas se estruturaram historicamente de modo simultâneo, mas independentemente (Hull, 1975). O principal fator para tal independência é o tipo de pergunta que cada uma faz: enquanto a genética de transmissão se ocupa do cruzamento de genes para estudar suas diferenças e sua distribuição na progênie, a genética molecular é um estudo de processos de biossíntese e da relação genótipo-fenótipo (Hull, 1975).

Quando ocorreu a troca terminológica dos “fatores” e “caracteres” mendelianos pelos respectivos termos “alelos/genótipo” e “fenótipo” da genética moderna, sem a fundamentação fisiológica da genética molecular, os adjetivos “dominante” e “recessivo” (originalmente referentes aos caracteres mendelianos) foram atribuídos como adjetivos para os fatores mendelianos em seu correspondente moderno, ou seja, passou-se a utilizar a expressão “alelo dominante/recessivo”. Porém, do ponto de vista fisiológico, não faz sentido denominar alelos como “dominantes” ou “recessivos”, pois a dominância e a recessividade são propriedades do fenótipo (Pasternak, 2002; Griffiths et al., 2011).

Esse desvio de nível da atribuição do fenômeno da dominância origina distorção conceitual, mas é a forma *status quo* de tratar genética no discurso didático. O equívoco está em associar que um alelo seja o responsável direto por um fenótipo final observável, sem considerar inúmeros outros fatores envolvidos. Correlação não é o mesmo que causa: identificar que um alelo tem relação com a produção de um fenótipo não necessariamente quer dizer que ele é o gene responsável pela proteína final que representa aquele fenótipo. Para tornar mais evidente como é inadequada a abordagem tradicional que se tem feito do discurso didático de genética, podemos explorar alguns casos clássicos mendelianos.

Sabe-se, por exemplo, que o fenótipo amarelo é o dominante na cor das ervilhas, porém todas as ervilhas, independentemente de seu genótipo, iniciam em um estado imaturo de pigmentação verde. Ao amadurecer, se

tornam amarelas a partir da degradação da clorofila por meio da ação sequencial de uma série de enzimas. O gene "SGR" sequencia uma das enzimas envolvidas no processo da degradação (Offner, 2011).

Em indivíduos nos quais as sementes permanecem verdes (fenótipo recessivo), a proteína SGR mutada, que é produzida a partir dos alelos na versão "v", não é funcional (porque ocorre nela a adição de dois aminoácidos: leucina e isoleucina). Nessas plantas (vv), a degradação da clorofila não acontece e a cor verde permanece mesmo na semente madura (efeito "stay green") (Garcia e Pereira, 2016).

Portanto, nas plantas homozigóticas amarelas (VV), a enzima SGR é funcional, enquanto nas plantas homozigóticas verdes (vv), a enzima SGR não é funcional e a clorofila não é degradada. Em ambos os casos, não ocorre o fenômeno da dominância, afinal, dominância é um fenótipo que só emerge no heterozigoto (Griffiths et al., 2011). Assim, ainda cabe a pergunta a ser respondida: Por que as ervilhas do indivíduo heterozigoto (Vv) têm fenótipo amarelo? Ou seja, por que a cor verde não é detectável?

Em uma sentença de genética molecular, podemos dizer que a planta heterozigota (Vv) terá metade da quantidade de enzima funcional produzida em um homozigoto (VV). Ainda assim, as enzimas estarão em quantidade suficiente para degradar toda a clorofila das ervilhas, tornando-as amarelas quando maduras (Thomas et al., 1996). O modo como o genótipo se relaciona com o fenótipo no caso da cor das ervilhas mendelianas ilustra como hipóteses simplificadas podem estar equivocadas (como, por exemplo, imaginar que o pigmento verde é resultante da expressão gênica do alelo "v" diretamente).

Assim, uma hipótese de que cada alelo originaria um pigmento diretamente a partir da síntese proteica estaria equivocada para ambas as cores, uma vez que os alelos "V" e "v" estão relacionados com a produção de uma dentre as várias enzimas que têm função na degradação da clorofila.

Para compreender este e outros casos de dominância, é possível fazer uma interpretação auxiliada pelo conceito de "haplossuficiência", que significa que em uma célula contendo apenas uma única cópia de um gene (condição heterozigótica), haverá fornecimento de produto gênico suficiente (em geral, uma proteína) para a realização dos processos celulares normais: "Em um heterozigoto, por exemplo Aa, no qual 'A' é o alelo selvagem, a única cópia do alelo A é suficiente para produzir produto proteico que desempenhe a função normal." (Griffiths et al., 2011, p. 46).

A haplossuficiência também pode ser constatada na cor do fruto da ervilha (vagem) de modo inverso ao observado na cor do cotilédone da semente, uma vez que o fenótipo verde é o que se concretiza em uma planta heterozigota (genótipo "Gp.gp"). Ou seja, na vagem, a cor amarela é recessiva e a cor verde é dominante. Enquanto na cor da ervilha a forma mutante representa uma falha na degradação da clorofila (resultando em fenótipo verde nos indivíduos "vv"), a forma mutante relacionada à cor da vagem falha em desenvolver cloroplastos normais nos frutos em indivíduos "gp.gp" (originando o fenótipo amarelo) (ver Ellis et al., 2011; Smýkal, 2014).

O caractere mendeliano da textura da ervilha (lisa/rugosa) também já foi estudado no que diz respeito aos aspectos moleculares e fisiológicos. Nesse exemplo, o fenótipo dominante é a textura lisa e o fenótipo recessivo é o rugoso. O alelo "R" (relacionado ao fenótipo dominante) codifica uma enzima que converte uma forma não ramificada de amido a uma forma ramificada. Ao contrário, o alelo mutado "r" (relacionado ao fenótipo recessivo) codifica uma forma disfuncional dessa enzima, o que causa acúmulo de amido não ramificado (Reece et al., 2015). A enzima disfuncional é codificada por um alelo que exhibe uma mutação caracterizada por uma inserção de uma sequência nucleotídica (Reid e Ross, 2011) - a diferença se deve a 800 pares de bases inseridos no alelo "r" (Heim, 1991).

Na planta homocigota "rr", o armazenamento de amido não ramificado causa a entrada de água em excesso na semente, por osmose. A semente posteriormente "secará", originando o enrugamento (fenótipo rugoso). Porém, em uma planta heterocigota (Rr), há produção de enzima suficiente para converter o amido, de modo que não ocorra a entrada excessiva de água, nem o posterior enrugamento. O fenótipo da planta heterocigota é, portanto, indistinguível do fenótipo da planta homocigota "RR", pois exhibe textura lisa nas sementes de ervilha (Reece et al., 2015).

Em nível bioquímico, pode-se dizer que ocorre codominância no heterocigoto, pois ambos os alelos originam produtos proteicos. Porém, ao nível do organismo, a verificação fenotípica permite a identificação de dominância completa entre os fenótipos liso e rugoso no heterocigoto (o que ratifica o caráter fenomenológico da dominância).

Em outras palavras, para determinar o tipo de dominância, qualquer fenótipo que esteja sob análise deve ser verificado em diferentes níveis de fenótipo, pois a relação de dominância e recessividade pode variar do nível molecular ao nível do organismo (Hull, 1975). Ademais, cada caso de dominância requer elucidação específica, sem generalizações, pois podem incluir, por exemplo, mecanismos causais de perda do RNAm no *splicing* (Schnur, 1996), defeito na estrutura tridimensional (dobras) da enzima, retenção em organelas citoplasmáticas ou outros (ver Ceschim e Caldeira, 2020).

O fenômeno da dominância, portanto, requer explicações fisiológicas que tendem a ser particulares, não generalizáveis que, muitas vezes, ainda estão sob investigação científica. Se estudantes não dispõem de nenhum exemplo para analisar a fisiologia subjacente à dominância, é esperado que produzam hipóteses intuitivas alternativas, incorretas.

## *2.2 Dificuldades na aprendizagem de dominância*

Conforme a literatura aponta, estudantes tendem a fazer uso dos termos "dominante" e "recessivo" sem entender por que um alelo selvagem e um alelo mutante originariam fenótipos recessivos/dominantes (Seipelt-Thiemann et al., 2012).

Algumas razões poderiam explicar a dificuldade para a aprendizagem de genética, tais como (Marbach-Ad e Stavy, 2000): (1) Níveis microscópicos são mais difíceis de entender do que o nível macroscópico, pois são geralmente ensinados de maneira teórica ou abstrata (2) Explicações de nível microscópico normalmente não podem ser extrapoladas

(erroneamente) para níveis macroscópicos. (3) Um nível (por exemplo, o nível macroscópico) pode constituir uma disciplina, e o outro nível (por exemplo, nível molecular) constitui diferentes disciplinas. (4) Vários níveis de organização devem ser integrados a fim de compreender os processos subjacentes aos fenômenos genéticos e a hereditariedade. Pensar simultaneamente uma variedade de conceitos e processos em diferentes níveis de organização pode dificultar a visão de cada conceito como parte de um todo inter-relacionado.

Ao concluir um processo formativo, por exemplo, espera-se que um estudante disponha de uma compreensão da genética clássica e de uma compreensão da genética molecular, mas não se espera que ele seja capaz de pensar acerca de como as duas disciplinas estão relacionadas (porque ele nem foi solicitado a fazer esse exercício de pensamento durante a formação) (Pavlova e Kreher, 2013).

Além das questões epistemológicas supracitadas que caracterizam a dificuldade para compreender o conhecimento relativo à hereditariedade, também existe um agravante importante registrado na literatura, que é a linguagem didática (Ceschim, Ganiko-Dutra e Caldeira, 2020). Foi observado por exemplo, que expressões como "O gene para olho vermelho", "O gene vermelho", ou "O gene letal", tendem a estimular um entendimento de que genes e traços constituem uma mesma unidade (Thörne, Gericke e Hagberg, 2013). Tal fusão gene-traço seria um dificultador para a aprendizagem da relação genótipo-fenótipo, pois se são entendidos como a mesma coisa, não seria necessário descrever a relação entre eles.

No que tange especificamente à dominância e recessividade, pesquisadores (Thörne, Gericke e Hagberg, 2013) identificaram no discurso de professores de Biologia a descrição de "genes dominantes" (sic!) como entidades atuantes que vencem/dominam os "genes recessivos" (sic!). Esse tipo de explicação pode estimular estudantes a entender que existe uma interação "supressora" entre alelos.

Como foi possível exemplificar em um estudo (Mahadeva e Randerson, 1982), outros equívocos comuns são: (a) acreditar que a dominância e a recessividade sejam propriedades fixas dos genes; (b) entender que genes dominantes são genes "bons" e os recessivos são "ruins"; (c) os genes dominantes na condição heterozigótica são entendidos como necessários para mascarar os efeitos deletérios de genes recessivos "ruins"; e (d) como as mutações são inerentemente ruins, elas precisam ser recessivas, pois se fossem dominantes, se propagariam rapidamente, produzindo efeitos desastrosos na população, talvez até a extinção.

Tais pressupostos estão fundamentados na aceitação de uma ideia equivocada acerca da dominância: a crença de que as características dominantes são mais numerosas em uma distribuição populacional (ver Allchin, 2000; Smith, 2014).

Lewin (2000) indica que uma explicação alternativa fornecida por estudantes seria a de que a recessividade resulta da ausência de uma característica, produto gênico ou proteína (também equivocada).

Em consonância, Redfield (2012) descreve o que a autora denomina como “problema da dominância”, com o apontamento dos seguintes equívocos: (1) a maioria dos estudantes acredita erroneamente que os alelos são intrinsecamente dominantes ou recessivos; (2) estudantes de nível superior que cursam genética não têm explicações para justificar como um alelo dominante (sic!) se comporta com relação ao outro – muitas vezes nem entendem a pergunta: “Por que um alelo se comporta como dominante (sic!) com relação ao recessivo?”; (3) quando os estudantes respondem alguma hipótese, normalmente apontam que alelos dominantes (sic!) devem ativamente “desligar” seus parceiros recessivos, talvez atuando como repressores ou por meio de mecanismos epigenéticos.

Tais incompreensões acerca da dominância são influenciadas pelo caráter metafórico que está envolvido no uso da palavra “dominante”, uma vez que significados implícitos podem afetar a estruturação de diversos conceitos relacionados ao seu uso (Allchin, 2000; Pearson e Hughes, 1988).

Considerando as conclusões das pesquisas supracitadas, faz-se necessária a presente investigação, uma vez que incompreensões referentes à dominância foram constatadas por outros pesquisadores (confirmando a hipótese do trabalho). Assim, torna-se relevante o aprofundamento da discussão acerca das características de tais incompreensões (como as premissas que fundamentam as hipóteses dos estudantes).

### **3. Metodologia**

Com o objetivo de determinar quais são as explicações atribuídas por professores de biologia em formação para o fenômeno da dominância completa, um questionário escrito (Flick, 2012) foi aplicado e foram obtidas respostas de 101 participantes. A coleta por meio de questionário apresenta um caráter padronizado devido à formulação prévia e unificada das perguntas, bem como do sequenciamento. Com as perguntas designadas de forma idêntica a todos os participantes, o questionário escrito permite a obtenção de respostas comparáveis (Flick, 2012).

Os participantes estavam no curso de Ciências Biológicas na modalidade de licenciatura (ou seja, a modalidade destinada à formação de professores) de dois campi da Universidade Estadual Paulista “Júlio de Mesquita Filho” localizados no estado de São Paulo, Brasil. Todos os participantes já tinham cursado a disciplina de genética, o que permitiria que fizéssemos questionamentos referentes aos conteúdos de ensino de nível superior referentes à dominância.

#### *3.1 Procedimentos da coleta de dados*

O questionário foi constituído por duas perguntas “de conhecimento” (ver Sampieri, Collado e Lucio, 2013), ou seja, são questões direcionadas para a obtenção de indicadores de compreensões dos indivíduos.

Inicialmente, uma primeira versão foi aplicada para um n=26 na data de 23 de outubro de 2017. A primeira pergunta explicava os dados da genética de transmissão das ervilhas mendelianas no que se refere ao fenótipo de textura lisa/rugosa da semente. A primeira versão do questionário foi assim constituída:

Pergunta 1: Considere as ervilhas de Mendel que exibem textura lisa ou rugosa. Houve constatação de que a ervilha rugosa é manifestada somente em condição homocigota recessiva (rr), enquanto a condição lisa se manifesta com genótipo homocigoto dominante (RR) ou heterocigoto (Rr). Por que no indivíduo heterocigoto a característica "textura lisa" se manifesta?

Pergunta 2: Quando um indivíduo possui um genótipo heterocigoto e exibe a característica dominante, a característica recessiva não se manifesta. O que isso significa na fisiologia? [em outras palavras, como a biologia funcional conseguiria explicar a manifestação de um fenótipo e o impedimento da manifestação de outro fenótipo?]

As duas perguntas se referem ao mesmo fenômeno e fazem o mesmo questionamento. Porém, a primeira pode permitir que o estudante ignore a necessidade de se atribuir uma explicação fisiológica, o que já será um primeiro teste, a saber: os participantes atribuem explicação fisiológica para a primeira pergunta? Como é um de nossos objetivos criar um repertório de explicações fisiológicas que os participantes atribuem para a dominância, foi proposta também a segunda pergunta, solicitando explicitamente que o respondente se posicione acerca dos aspectos fisiológicos envolvidos. Os participantes só poderiam acessar a segunda pergunta após responderem à primeira.

Diante dos dados obtidos nessa primeira aplicação do questionário, a pergunta foi alterada para as próximas coletas. Tendo em vista que os participantes não conseguiram atribuir explicação fisiológica para o exemplo das ervilhas lisas/rugosas na primeira coleta supracitada, decidimos escolher um exemplo mais conhecido: o fenótipo das sementes de ervilhas amarelas/verdes. A nova pergunta utilizada nos novos questionários foi a seguinte:

Pergunta 1: Considere as ervilhas de Mendel que exibem cor amarela ou verde. Houve constatação de que a cor verde é manifestada somente em condição homocigota (vv), enquanto a cor amarela se manifesta com genótipo homocigoto (VV) ou heterocigoto (Vv). Por que no indivíduo heterocigoto a característica "amarela" se manifesta?

A Pergunta 2 utilizada foi idêntica à do questionário anterior. Várias coletas foram realizadas: as primeiras nos dias 18 de junho de 2018 e 21 de junho de 2018, posteriormente em 27 de junho de 2018 e as demais nos dias 05 e 09 de julho de 2018, 18 de julho de 2018, 17 de outubro de 2018, 15 de abril de 2019 e 02 de março de 2020.

Como parâmetro de resposta para a Pergunta 1, esperava-se a indicação de que a dominância ocorre devido à síntese de produtos diferentes por cada alelo, de modo que no indivíduo heterocigoto ocorrerá a manifestação detectável somente do fenótipo dominante decorrente do efeito do produto originado por seu respectivo alelo. Os processos envolvidos na perda/manutenção da cor da clorofila nas sementes de ervilha também eram esperados como resposta.

No que diz respeito à Pergunta 2, esperava-se a indicação de mecanismos genéricos para exemplificar, como a produção de enzimas

disfuncionais que causa no indivíduo heterozigoto a perda do efeito do alelo relacionado ao fenótipo recessivo.

Os participantes concordaram com o termo de consentimento livre e esclarecido e tiveram suas identidades protegidas, sendo identificados como "A\_" ou "B\_". A primeira coleta referente às ervilhas rugosas/lisas tem os participantes representados pela letra "B" e a segunda coleta referente às ervilhas amarelas/verdes tem os participantes representados pela letra "A".

### *3.2 Procedimentos da análise dos dados*

A análise das respostas discursivas obtidas foi concretizada por meio de dois procedimentos: a Análise Textual e a Análise de Conteúdo.

Para determinar quais palavras foram mais utilizadas e quais associações de palavras são mais frequentes nas respostas, o procedimento empregado foi a Análise Textual (Nascimento-Schulze e Camargo, 2000) viabilizado pelo *software* estatístico IRaMuTeQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*) que foi criado por Pierre Ratinaud em 2009. O programa funciona a partir do *software* R (versão 4.0.3), compatível com a linguagem Python. O *software* IRaMuTeQ é licenciado pelo *Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales* (LERASS) da Universidade de Toulouse na França.

Para as análises do banco de dados, que é elaborado em formato de texto (denominado "corpus textual"), foi realizada a verificação estatística da ocorrência de palavras. A *Análise de Similitude* empregada permite identificar como foi realizada a estruturação do discurso, de modo que a distribuição de palavras relacionadas é destacada por grau de relevância com base na frequência em que aparecem as coocorrências (palavras que ocorrem de modo simultâneo nos trechos).

O resultado da Análise de Similitude pode ser expresso no formato de "árvores máximas", que são representações das coocorrências, nas quais são originados os "halos", agrupamentos de elementos que se relacionam nas frases. Assim, as árvores máximas indicam relação entre os itens e a importância de cada relação, uma vez que os ramos que ligam halos e que ligam termos têm o tamanho da espessura proporcional à relação entre as unidades. Do mesmo modo, o tamanho das palavras é proporcional à frequência de aparecimento no banco de dados. As palavras mais frequentes na maioria das frases originam os elementos centrais dos halos.

Alguns termos foram ligados para que a interpretação do *software* fosse única, como: "textura\_lisa", "alelo\_dominante", "alelo\_recessivo", "característica\_lisa", "característica\_rugosa", "cor\_verde", "cor\_amarela", "gene\_recessivo", "gene\_dominante", "característica\_amarela", "característica\_recessiva" e "característica\_dominante".

Palavras com o mesmo radical foram substituídas por uma única forma (por exemplo: as palavras "recessivo", "recessivos", "recessiva" e "recessivas" foram unificadas na forma "recessivo"). Esse procedimento de unificação por meio de radical idêntico (lematização) também foi realizado com os termos: "manifestar", "expressar", "inibir", "sobressair", "impedir", "bloquear", "dominar", "suprimir", "silenciar" e "determinar" no corpus textual, utilizando o recurso "localizar e substituir" do editor de texto.

Respostas em branco e respostas que indicavam que o graduando não sabia responder (como “não sei”, “não lembro” e similares) não constituíram o banco de dados que seria utilizado. A partir das análises realizadas pelo IRaMuTeQ, uma figura de árvore máxima foi gerada como resultado da *Análise de Similitude* e também foi gerada uma figura tipo “nuvem de palavras” para indicar uma relação de proporcionalidade direta entre o tamanho das palavras e sua frequência no corpus textual.

Com o objetivo de identificar as particularidades das explicações que os graduandos estruturaram em suas respostas, os dados também foram verificados por meio da Análise de Conteúdo – segundo a proposta de Laurence Bardin (2011). O procedimento consiste em uma análise categorial, por meio da classificação de um texto segundo frequências de presença e ausência de um item. As categorias foram elaboradas *a posteriori*, por meio do agrupamento de elementos particulares que, após serem reunidos por critérios de semelhança, puderam receber um título que nomeava uma categoria formada. As unidades de registro utilizadas para discriminar trechos que foram classificados em cada categoria constam nos resultados.

Devido à questão de limitação de espaço, não será possível descrever todo o processo categorial, mas vamos exemplificar alguns dos passos. Na leitura flutuante seguida por uma leitura mais aprofundada, já foi possível identificar que várias respostas não tratavam de fisiologia (contrariando o que era esperado para o questionamento). Esse conjunto de respostas que “não trata de fisiologia” se tornou antagônico do conjunto que “trata de fisiologia”, fazendo emergir duas grandes categorias.

Para classificar respostas nas categorias, indicadores foram listados. Por exemplo, para agrupar as respostas nas quais os participantes citavam a produção biossintética (transcrição/tradução), os indicadores foram “expressar”, “proteínas”, “decodificar”, “transcrever” e “enzima”.

## **4 Resultados**

### *4.1 Análise Textual*

Como resultado da Análise de Similitude realizada por meio do *software* IRaMuTeQ, a árvore máxima da Figura 1 foi gerada. No núcleo central da árvore está a palavra “manifestar” da qual partem ramificações. Os ramos com maiores graus de conexidade com o núcleo são: manifestar/característica e manifestar/dominante. A árvore máxima subdivide-se em quatro halos, a saber: (a) “expressar”, (b) “manifestar”, (c) “dominante” e (d) “heterozigoto”.

No halo de “manifestar” é possível observar as conexões manifestar/característica, manifestar/característica\_dominante e manifestar/apenas. Também ocorrem as conexões característica\_recessiva/impedir e presença/gene\_dominante.

No halo “expressar”, uma relação frequente foi entre expressar/característica. O termo “expressar” também aparece associado ao termo “não”, que por sua vez está relacionado aos termos “proteína”, e “produzir”.

No halo “dominante”, há uma conexão frequente com alelo/dominante, outra com dominante/característica\_amarela, outra com dominante/gene, outra com dominante/cor\_amarela/cor\_verde e dominante/cor. Relações menos frequentes incluem dominante/sobressair e dominante/recessivo.

No halo “heterozigoto”, há a conexão mais frequente heterozigoto/indivíduo e termos que indicam a relação de presença (“presente”, “apresentar” e “possuir”).

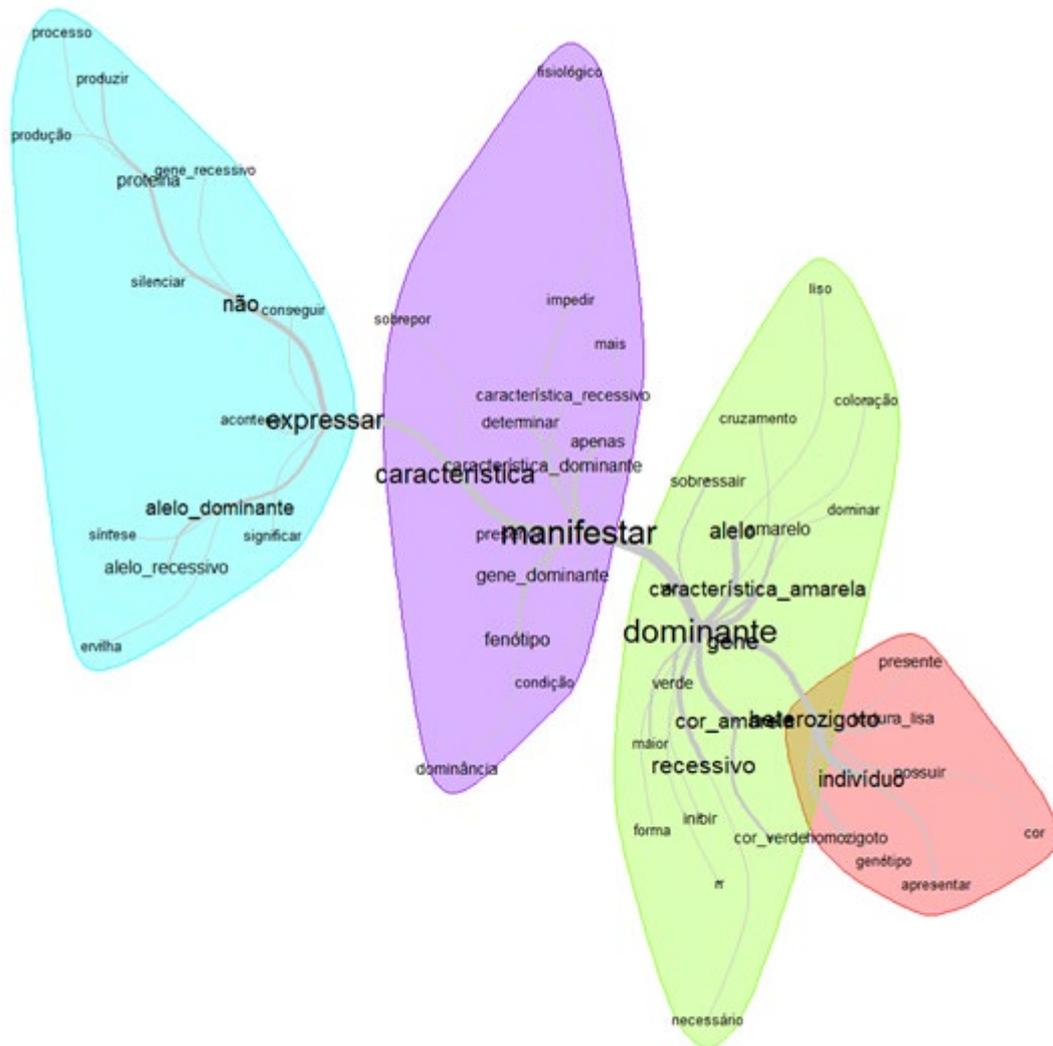


Figura 1. Análise de similitude das respostas dos graduandos do curso de Ciências Biológicas. Elaborada pelas autoras. Extraído do software IRaMuTeQ.

Alguns segmentos de texto ilustram tais relações estabelecidas:

A10 “Pois a característica “amarela” possui dominância sobre a característica verde, ou seja, a amarela é dominante enquanto a verde é recessiva. Em heterozigoto sempre a característica dominante se manifesta.”

A11 “A característica amarela se manifesta, pois, o indivíduo heterozigoto possui gene (Vv), e segundo os estudos de Mendel, o gene (V) apresenta-se como dominante, ou seja, se manifesta suprimindo a manifestação do gene (v). Dessa forma os indivíduos heterozigotos expressam a cor amarela.”

A “nuvem de palavras” (Figura 2), também gerada como resultado de uma análise realizada pelo software IRaMuTeQ, indica o uso mais frequente do termo “manifestar” (84 vezes), seguido pelos termos: “dominante” (69), “expressar” (62), “característica” (51), “alelo” (47), “gene” (47), “não” (40), “recessivo” (40), “heterozigoto” (39), “cor\_amarela” (35), “indivíduo” (34), “alelo\_dominante” (30), “característica\_amarela” (29), “proteína” (25), “fenótipo” (25), o genótipo “vv” (23) e outros.



Figura 2. “Nuvem de palavras” que representa o uso de termos mais/menos frequente nas respostas dos graduandos do curso de Ciências Biológicas. Elaborada pelas autoras. Extraído do software IRaMuTeQ.

Essas palavras e expressões podem ser tomadas como um conjunto que constitui o vocabulário mais comumente empregado para estruturar as respostas dos participantes.

#### 4.2 Análise de Conteúdo

##### 4.2.1 Elaboração de categorias

Durante a leitura dos dados, foram elaboradas as categorias *a posteriori* e as respostas foram classificadas com base na semelhança e repetição de unidades de registro.

Nas respostas da Pergunta 1, identificamos a emergência de dois padrões: os participantes (1) apenas reafirmavam os dados do enunciado ou (2) tentavam atribuir uma explicação fisiológica. Assim, formaram-se

categorias antagônicas. Ademais, "reafirmar os dados do enunciado" significa que os respondentes não extrapolavam afirmações acerca da genética de transmissão, ou seja, não tratavam de genética molecular e não tratavam de fisiologia. Para que fosse possível classificar respostas na segunda categoria, seriam necessários indicadores de que o participante estava fazendo afirmações relativas à dimensão fisiológica para se diferenciar da mera "reafirmação de dados de genética de transmissão".

Tais indicadores seriam "mecanismos fisiológicos" (de forma metafórica ou não). Os indicadores que obtivemos foram: "expressar", "inibir"/"impedir", "silenciar", "sobressair", "sobreviver", "prevalecer", "produzir"/"sintetizar", "recobrir" e "dominar". Os verbos foram aqui indicados no infinitivo, mas apareceram também em outras flexões nas respostas produzidas pelos participantes.

Quanto à categoria restrita à genética de transmissão, outras subcategorias foram criadas. Alguns utilizaram a ideia da "presença", que corresponde a dizer que "o alelo/característica dominante está presente no indivíduo" como explicação. Os indicadores foram "presente/presença", "ter", "possuir", "haver" e respectivas flexões.

Mais uma subcategoria criada foi referente à haplossuficiência do alelo dominante (sic!), ou seja, a explicação afirmava que um único alelo dominante (sic!) determinaria o fenótipo. Os indicadores incluem "apenas", "um", "único" e "sozinho".

Dentre as explicações classificadas na primeira categoria, a maioria das respostas que restaram possui um elemento comum: a relação de dominação. Neste caso, os participantes indicam que um alelo é dominante em relação a outro, ou então que uma característica é dominante em relação à outra. Como indicadores, utilizamos os termos "dominância/dominante" ou "recessivo" que estivessem em uma frase que contivesse tal relação causal.

Quanto às respostas que apresentavam explicação fisiológica, três subcategorias foram delineadas: a primeira representada por explicações metafóricas referentes a interações entre alelos (categoria 2.1); a segunda representada por explicações de que dominância se deve a alterações qualitativas na expressão gênica (produção de proteínas/RNA) – categoria 2.2; e a terceira representada por argumentos referentes à dominância ser resultante de diferenças na *quantidade* de expressão gênica (categoria 2.3) – cujos indicadores utilizados foram de intensidade: "maior" ou "mais".

Sendo assim, foi possível organizar seis subcategorias que foram reagrupadas em duas categorias mais abrangentes, a saber: "1 Argumentos restritos à genética de transmissão" e "2 Argumentos com explicação fisiológica". As respectivas subcategorias são: "1.1 Presença do alelo dominante"; "1.2 Haplossuficiência"; "1.3 Dominância do alelo"; e "2.1 Ação metafórica do alelo/gene dominante (sic!) sobre o alelo/gene recessivo (sic!)"; "2.2 Alteração na produção de proteínas/RNA"; e "2.3 Alteração na quantidade de produção de proteínas/RNA".

Na análise da Pergunta 2, foi possível identificar várias respostas "não sei" e algumas respostas em branco, o que não ocorreu para a Pergunta 1. Algumas explicações exibiam uma estrutura que sugere um caráter de

hipótese, e não uma convicção. Assim como aconteceu nas respostas à pergunta anterior, alguns graduandos citaram somente dados referentes à genética de transmissão, sem explicação fisiológica. Porém, não foi necessário classificar o tipo de explicação restrita à genética de transmissão, pois as afirmações identificadas são genéricas e não permitiam a construção de subcategorias.

Alguns utilizaram linguagem metafórica para fazer referência aos mecanismos fisiológicos e outros citaram transcrição e síntese proteica em suas explicações (de modo semelhante aos resultados da Pergunta 1).

A concepção alternativa sobre o fenótipo dominante ser o mais frequente na distribuição populacional apareceu em algumas respostas. Assim, além das categorias pré-existentes originadas da classificação das respostas da Pergunta 1, que foram reaproveitadas em decorrência da semelhança dos resultados, foi possível criar a categoria de "Maior ocorrência/frequência na população", por meio de indicadores como "probabilidade", "ocorrência" e "frequência".

Também foi criada a categoria "Dominante é mais forte e o recessivo é o mais fraco" por meio de indicadores, como "forte" e "fraco". Por fim, a categoria "O fenótipo resulta das interações com a influência do meio" foi criada a partir dos indicadores "ambiente" e "meio". As categorias formadas a partir do procedimento supracitado, portanto, são: "1) Argumentos restritos à genética de transmissão"; "2) Argumentos com explicação fisiológica"; "3) Maior ocorrência/frequência na população"; "4) Dominante é mais forte e o recessivo é o mais fraco"; "5) O fenótipo resulta das interações com a influência do meio"; "6) Não respondeu"; e "7) Não sabe".

As subcategorias da segunda categoria foram: "2.1) Ação metafórica do alelo/gene dominante (sic!) sobre o alelo/gene recessivo (sic!)"; "2.2) Alteração na produção de proteínas/RNA"; e "2.3) Alteração na quantidade de produção de proteínas/RNA".

#### *4.2.2 Análise das frequências das respostas e apresentação de exemplos*

Quanto à Pergunta 1, as respostas da primeira categoria contemplam explicações que tratam da distribuição de alelos, sem fazer menção a mecanismos moleculares e fisiológicos. A primeira subcategoria (1.1) significa que o participante explica a dominância justificando que a presença/posse de um alelo ou característica dominante causa o aparecimento do fenótipo dominante. No total, 16 respostas foram classificadas nessa categoria. Alguns exemplos de respostas da categoria 1.1 são:

A29: Pois a cor amarela *está presente* no gene V.

B12: Porque o indivíduo heterozigoto *possui* um alelo dominante B.

A segunda categoria (1.2) indica que o participante explica a dominância apontando que a presença de apenas um alelo é suficiente para causar o fenótipo dominante (haplossuficiência), como é possível observar na resposta a seguir: A18 "A cor amarela se manifesta mesmo no indivíduo heterozigoto pois o alelo V, *sozinho*, é capaz de determinar a característica fenotípica da cor amarela". As respostas classificadas nessa categoria totalizam 12 graduandos.

A terceira categoria indica que o participante explicou o fenômeno apontando a dominância de um alelo sobre o outro. Foi a categoria com maior número de respostas (71 unidades). Alguns exemplos de respostas para a categoria 1.3 incluem:

A3: Porque a característica amarela *é dominante em relação à verde*, que é recessiva.

A10: Pois a característica "amarela" possui *dominância sobre a característica verde*, ou seja, a amarela é dominante enquanto a verde é recessiva. Em heterozigoto sempre a característica dominante se manifesta.

Na categoria 2.1, as respostas tratam metaforicamente de meios pelos quais a dominância ocorre e nela houve a ocorrência de 23 repostas. Os verbos que foram inseridos nessa categoria foram "inibir", "sobressair", "prevalecer", "suprimir", "silenciar", "recobrir", "expressar", "sobrepor", "impedir" e "sobreviver". Alguns exemplos de respostas são:

A36: Pois o gene responsável pela manifestação da cor amarela é considerado dominante em relação ao gene da expressão da cor verde, se comportando como *inibidor da característica do outro*.

A58: Porque a característica "amarela" é dada pelo gene dominante, de forma que *este silencia o gene diferente* presente no indivíduo.

Na categoria 2.2, quatro respostas tratam de explicações referentes à produção de proteína/RNA envolvida na expressão gênica dos alelos dominantes. Alguns exemplos de respostas incluem:

Na categoria 2.3, dois participantes tratam a dominância como resultante de diferenças entre quantidades de proteínas/RNA entre o alelo dominante e o recessivo (sic!). Constam as duas respostas abaixo:

A2: [...] A cor amarelada implica na síntese de proteínas que colorem a ervilha assim como a verde. Porém, a condição amarela se faz fruto de alelo dominante em relação a verde, logo a proporção de RNAm traduzido deste alelo dominante *é muito mais* do que o RNAm do verde, e assim ela se manifesta pela cor amarela.

A78: Pois o alelo dominante (V) se expressa com maior intensidade que o recessivo (v), então a característica amarela é expressa pelo alelo dominante.

Para fazer a classificação das respostas nas referidas categorias, em alguns casos, uma mesma frase foi desmembrada em partes, porque um participante pode ter utilizado mais de uma estratégia explicativa. Por exemplo, o aluno A1 utiliza "1.3 Dominância do alelo/característica" e 2.2 "Ação metafórica do alelo/gene dominante (sic!) sobre o alelo/gene recessivo (sic!)", como é possível observar a seguir: "Porque o *gene que define a cor amarela é dominante* [1.3] e *se sobressai em cima do gene recessivo*. [2.2]"

A comparação das frequências de respostas classificadas em cada categoria pode ser observada na Figura 3. Na representação da figura, estão indicadas quantas vezes indicadores de uma mesma categoria foram identificados e classificados nessa mesma categoria.

Por exemplo, a categoria 1.3 teve 71 trechos de respostas contendo indicadores passíveis de agrupamento por semelhança. Nelas, os participantes explicam a dominância apontando que um alelo é dominante com relação ao outro. Foi a categoria mais frequente, representada por uma quantidade de respostas aproximadamente 208,7% superior à quantidade da segunda categoria mais frequente (categoria 2.1).

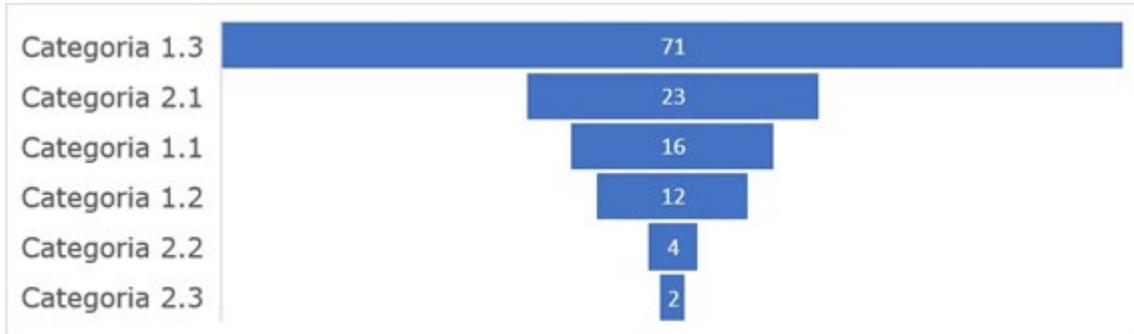


Figura 3. Gráfico das frequências de respostas de cada categoria referente à Pergunta 1. As categorias são 1.1 Presença do alelo dominante; 1.2 Haplosuficiência; 1.3 Dominância do alelo; 2.1 Ação metafórica do alelo/gene dominante (sic!) sobre o alelo/gene recessivo (sic!); 2.2 Alteração na produção de proteínas/RNA; 2.3 Alteração na quantidade de produção de proteínas/RNA. Elaborada pelas autoras.

No que se refere à Pergunta 2, apenas seis respostas foram consideradas dentro da primeira categoria, cuja explicação não inclui elementos fisiológicos, apenas referência à genética de transmissão. As respostas dessa categoria fazem afirmações acerca da distribuição de alelos nos indivíduos, como se pode observar na fala a seguir: A8 "A característica recessiva não se manifesta pois o alelo não está presente nos dois pares de cromossomos homólogos."

Na segunda categoria, foram alocadas respostas nas quais os participantes atribuem algum tipo de explicação fisiológica para a dominância. Na subcategoria 2.1 estão agrupadas aquelas respostas nas quais se faz afirmações metafóricas acerca da fisiologia da dominância e totalizam 19 respostas. Os termos utilizados foram: "silenciar", "dominar", "inibir", "suprimir", "sobressair", "bloquear" e "impedir" (com suas respectivas flexões). Na "nuvem de palavras" da Figura 4, constam as frequências representadas em tamanho maior/menor de cada termo que apareceu nas respostas classificadas nessa categoria: "bloquear" (apareceu uma vez), "dominar" (uma vez), "impedir" (duas vezes), "inibir" (seis vezes), "silenciar" (quatro vezes), "sobressair" (três vezes) e "suprimir" (três vezes):



Figura 4. "Nuvem de palavras" que representa a frequência do uso de cada termo metafórico nas respostas da Pergunta 2 que foram classificadas na categoria 2.1. Elaborada pelas autoras.

Alguns dos termos supracitados têm sentidos semelhantes, como "bloquear", "inibir", "silenciar", "impedir" e "suprimir". Do mesmo modo, "dominar" e "sobressair" têm sentidos semelhantes entre si. Em alguns casos, os participantes utilizavam como sujeito da ação o fenótipo (A26), em outros casos, o gene (A59): Exemplos: A26 "*A característica dominante inibe a recessiva, mas não a elimina*" e A59 "*A característica recessiva é inibida pelo gene dominante, por isso não é manifestada.*"

Em algumas respostas, os estudantes apontaram simplesmente que o alelo recessivo (sic!) está silenciado/suprimido/bloqueado/impedido (ver A5, a seguir). Em outros casos, indicaram um "sujeito" na frase para dizer quem silenciou/suprimiu/bloqueou/impediu: o gene/fenótipo dominante (ver B24, a seguir). As falas supracitadas são as seguintes: A5 "*As proteínas produzidas pelo gene dominante são as que vão determinar a característica. O gene recessivo fica silenciado, não produzindo proteínas.*" e B24 "*Significa que o alelo dominante inibe a expressão do recessivo.*"

Na categoria 2.2, os participantes explicam que a dominância resulta de alterações no processo de transcrição e/ou tradução e totalizam 28 respostas. Alguns apontam que não ocorrem tais processos de produção biossintética, como se pode ver em A3, "*Que apesar de o indivíduo apresentar a informação genética contida em seu DNA, a mesma não é expressa (sem transcrição e tradução) [...]*". Um participante aponta que interações diretas entre os alelos são a causa para a dominância, como se pode observar em: A34 "*Acredito que as interações moleculares de um alelo dominante com a de um recessivo devem fazer com que somente o alelo dominante seja transcrito.*"

Dois participantes indicam outra relação causal, pois entendem que são as proteínas sintetizadas pelos alelos que causam a dominância (e não a interação direta entre alelos): Exemplos: A58 "*Os genes da característica dominante expressam proteínas que silenciam o gene da característica recessiva, que não consegue ser expresso.*" e A63 "*Do ponto de vista fisiológico as proteínas existentes nos genes do alelo dominante inativam a expressão do recessivo. Impedindo sua manifestação.*"

Nas respostas da categoria 2.3 "Alteração na quantidade de produção de proteínas/RNA", quatro participantes explicam a dominância apontando que ocorrem alterações na quantidade de produtos biossintéticos que são originados dos alelos em questão, conforme as falas abaixo indicam:

A2: [...] Primeiramente, pode significar que a molécula/proteína gerada pelo alelo recessivo será produzida *em proporção maior* do que a proteína dominante, logo há preferência de tradução pelo RNAm originado do dominante. [...]

A79: O gene dominante [...] *expressa mais* proteínas que o recessivo.

Uma especificidade foi observada no restante da resposta do graduando A79 ("O gene dominante tem mais compatibilidade para expressar proteínas

naquele do que o recessivo, então ele expressa mais proteínas que o recessivo, porque o dominante é muito melhor. Para ser dominante, tem que ter mais compatibilidade do que a recessiva, ou seja, o maquinário que vai produzir aquele fenótipo daquele gene tem mais afinidade química com o gene dominante do que com o recessivo.”).

Como se nota, a principal estratégia explicativa para o fenômeno da dominância para este participante é a diferença na quantidade de proteínas que são produzidas pelos diferentes alelos, portanto tal resposta foi classificada na categoria 2.3. A peculiaridade dessa resposta, porém, está nas causas que o estudante aponta para essa diferença de quantidade. Além de caracterizar o gene dominante (sic!) como “melhor”, o graduando tem a hipótese de que deve haver mais compatibilidade entre o gene dominante (sic!) e o “maquinário” que produz o fenótipo em comparação com o gene recessivo (sic!), que teria menor compatibilidade.

Na categoria 3 “Maior ocorrência/frequência na população”, seis participantes fundamentam-se na concepção equivocada de que fenótipos dominantes são mais numerosos na distribuição populacional para justificar a causa da dominância, como se pode ver a seguir: A76 “Devido ao fato de que os genes (dominante) *ocorrem com maior frequência* fenotípica.” e B7 “Por não ser algo *mais frequente*, é mais difícil de se expressar [...]” [se referindo ao fenótipo recessivo].

Uma peculiaridade que ocorreu no restante da resposta de B7 foi o apontamento metafórico que descreve como ocorreria a dominância: “O ‘V’ é a parte de fora da caneta e o ‘v’ é a parte interna da caneta. Você não consegue chegar no tubo de tinta quando é ‘Vv’”. Ou seja, para o participante, a recessividade está internamente envolvida pela dominância.

Na categoria 4 “Dominante é *mais forte* e o recessivo é o mais fraco”, três participantes atribuem a causa da dominância a uma suposição de que alelo/fenótipo dominante (sic!) é mais forte que o recessivo, como é possível ver em A56: “Isso é explicado pelo processo de dominância do fenótipo, sendo que o “*mais forte*” é o que se manifesta.”

Na categoria 5 “O fenótipo resulta das interações com a influência do meio”, dois participantes julgam que a dominância seja causada, ao menos em parte, por influências ambientais, como se pode observar em A41: “Acho que a manifestação tem a ver com as *características do ambiente*, determinada condição externa e essa interação.”

A categoria 6 inclui os 13 participantes que não responderam à Pergunta 2 e a categoria 7 inclui os 27 participantes que declararam não saber responder. Juntos, totalizam 40 participantes sem uma resposta, como é possível observar nas frequências comparadas da Figura 5.

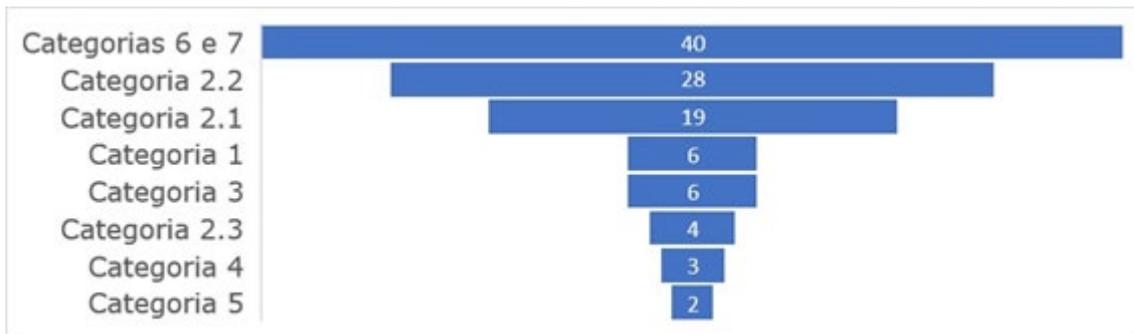


Figura 5. Gráfico das frequências de respostas de cada categoria referente à Pergunta 1. As categorias são: 1 Argumentos restritos à genética de transmissão; 2.1 Ação metafórica do alelo/gene dominante (sic!) sobre o alelo/gene recessivo (sic!); 2.2 Alteração na produção de proteínas/RNA; 2.3 Alteração na quantidade de produção de proteínas/RNA; 3 Maior ocorrência/frequência na população; 4 Dominante é mais forte e o recessivo é o mais fraco; 5 O fenótipo resulta das interações com a influência do meio; 6 Não respondeu; 7 Não sabe. Elaborada pelas autoras.

As categorias 6 e 7 juntas (que significam a ausência de uma resposta) representam o maior número de frequência ( $n=40$ ) dos resultados obtidos para a Pergunta 2. A segunda maior frequência ( $n=28$ ) foi da categoria 2.2, na qual os graduandos explicam que a dominância resulta de alterações na transcrição e/ou na tradução, mas não há indicações acerca de quais são esas alterações ou acerca de como elas ocorrem.

Assim, é possível afirmar obtivemos respostas que reafirmam dados do enunciado ou propõem hipóteses genéricas referente à dominância, especialmente quando a fisiologia envolvida no processo é solicitada (Pergunta 2) – questão para a qual alguns participantes não conseguiram oferecer uma explicação ou hipótese.

## 5. Discussão

Considerando o objetivo de indicar as explicações que os graduandos propõem para a dominância, foi possível determinar os elementos constitutivos do discurso utilizado pelos participantes.

A análise de frequência de palavras (nuvem de palavras) realizada com a utilização do IRaMuTeQ, por exemplo, listou os termos mais empregados para estruturar as explicações, tais como “manifestar”, “dominante”, “expressar”, “característica”, “alelo”, “gene”, “não”, “recessivo” e “heterozigoto”, o que indica os principais elementos mobilizados pelos professores em formação para explicar a dominância.

Por meio da *Análise de Similitude* resultante do IRaMuTeQ que estabeleceu as relações entre tais termos mais frequentes, podemos considerar esta como uma frase genérica que reúne os itens: ocorre a manifestação detectável de características dominantes em heterozigotos a partir da expressão de determinado alelo/gene, o que não ocorre no correlato recessivo.

Tendo determinado os elementos gerais da explicação que graduandos exibem, também é possível apontar elementos específicos, que foram produzidos a partir da Análise de Conteúdo. Algumas afirmações acerca dos

dados podem ser destacadas: os professores em formação (1) não identificam a necessidade de atribuir explicação fisiológica ao fenômeno da dominância e consideram resposta suficiente a repetição dos dados do enunciado, ainda que a fisiologia seja explicitamente solicitada; (2) admitem que interações diretas ocorrem entre os alelos; (3) explicam a recessividade do fenótipo apontando que o alelo não origina produtos, como Lewin (2000) já havia apontado; (4) indicam que o alelo recessivo é inativado por ação do alelo dominante – corroborando afirmações de Thörne, Gericke e Hagberg (2013) e Redfield (2012) – ou por ação dos produtos polipeptídicos gerados a partir dele; (5) apontam que o alelo dominante origina mais proteína em comparação com o recessivo; (6) apontam que a dominância ocorre no heterozigoto devido a sua distribuição prevalente na população – como também apontaram Allchin (2000) e Smith (2014); (7) caracterizam o alelo dominante (sic!) como mais forte ou como melhor que seu correlato recessivo – em consonância com o que apontam Mahadeva e Randerson (1982); (8) explicam que a dominância ocorre por influência do ambiente; além disso, (9) um graduando apontou que o elemento recessivo está “dentro” do elemento dominante, o que impede a detecção do recessivo; (10) um graduando explicou que o alelo dominante é mais compatível com o material orgânico e por isso produz mais produto polipeptídico; (11) um percentual importante (40%) de graduandos não respondeu somente a pergunta que solicitava a causa fisiológica da dominância, o que pode indicar a inabilidade para oferecer uma explicação, ainda que fosse hipotética e (12) nenhum graduando elaborou uma resposta correta para a segunda pergunta do questionário.

Quanto à linguagem, é possível afirmar que embora os adjetivos “dominante” e “recessivo” tenham sido originalmente propostos por Mendel com uma conotação fenomenológica, foram desdobrados em substantivos (“dominância” e “recessividade”) e em verbo (“dominar”), de modo que passaram a ter a equivocada conotação de processo fisiológico. O uso metafórico dos termos como adjetivos, ao ser ampliado para substantivos e verbos, passou a ser tomado como literal, originando estruturas frasais que admitem sujeitos (por exemplo: alelo, gene, fenótipo, característica) para o verbo “dominar”, como constatamos nas falas dos graduandos que participaram dessa pesquisa.

A compreensão equivocada de que o fenótipo dominante é mais numeroso nas populações também pode sofrer influência da conotação de preponderante, prevalente ou hegemônica que a dominância pode assumir no uso corriqueiro. Além disso, é possível associar as compreensões que indicam a superioridade do elemento dominante (quando é apontado como “melhor” ou “mais forte”) à conotação de poder/*status* que a dominância pode admitir também no uso cotidiano.

Por último, o termo “expressão/expressar”, muito utilizado no discurso didático durante as aulas, contém uma ambiguidade que pode contribuir com ruídos de comunicação. A expressão gênica, por exemplo, pode ser entendida como o processo de síntese de produtos a partir do material genético. Porém, a expressão do fenótipo, que seria sinônima de “manifestação do fenótipo”, se refere a produção de traços biológicos detectáveis (até mesmo macroscópicos). Sendo assim, o uso do termo “expressar” pode originar ambiguidade ou pode causar a unificação de

ambos os significados, estimulando a compreensão de que a expressão gênica é a própria expressão final do fenótipo.

Sendo assim, podemos inferir que a linguagem utilizada para conceber e comunicar os conceitos da dominância é fonte de incompreensões e de construção de interpretações não alinhadas com o conhecimento científico.

## **6. Conclusões**

Em consonância com outras pesquisas, nossos dados apontam várias concepções equivocadas e confirmam a hipótese de que existe uma deficiência formativa causada pela fragmentação entre genética de transmissão e a genética molecular. Vários de nossos resultados estão em conformidade com outras pesquisas, como por exemplo: (1) graduandos não têm explicações para responder por que um alelo se comporta como dominante com relação ao recessivo? – muitas vezes nem mesmo entendem tal pergunta (Redfield, 2012); (2) os estudantes empregam os termos “dominante” e “recessivo” sem saber explicar como os alelos originariam fenótipos (Seipelt-Thiemann, 2012); (3) são feitas extrapolações errôneas a partir da equiparação de uma explicação de nível macroscópico com uma explicação de nível microscópico (Marbach-Ad e Stavy, 2000); (4) a dominância e a recessividade são tomadas como propriedades dos alelos (Mahadeva e Randerson, 1982); (5) alelos dominantes (sic!) são entendidos como agentes que dominam/desligam os alelos recessivos (sic!) (Thörne, Gericke e Hagberg, 2013; Redfield, 2012); (6) alelos dominantes (sic!) são 'bons' e seus correlatos recessivos são "ruins" (Mahadeva e Randerson, 1982); (7) fenótipos dominantes são mais numerosos em uma distribuição populacional (ver Allchin, 2000; Smith, 2014); e (8) a recessividade é causada pela ausência de produto gênico ou proteína (Lewin, 2000).

Os participantes não atribuem explicações fisiológicas adequadas para a dominância completa, uma vez que alguns não conseguiram responder nada, outros reafirmaram os dados do enunciado da pergunta e outros exibiram uma pluralidade de compreensões equivocadas. Em conformidade com o objetivo inicial, este trabalho aprofundou as descrições das incompreensões referentes à dominância, ao documentar trechos e ao explorar semelhanças e particularidades dos dados.

Podemos destacar que, para além das correspondências das conclusões dessa pesquisa com as conclusões de pesquisas supracitadas (indicadas na discussão), nossa investigação permitiu a identificação (1) da compreensão equivocada de que o alelo dominante (sic!) produz mais quantidade de produto gênico em comparação com o recessivo; (2) da compreensão equivocada de que o maior número de fenótipos dominantes em uma distribuição populacional pode ser apontado como causa para a produção do fenótipo dominante; (3) de que a relação genótipo-ambiente é apontada como causa para a dominância; (4) da ideia de que a recessividade tem um elemento que localiza-se internamente ao elemento dominante (inviabilizando sua detecção).

Os trechos de frases dos participantes que foram inseridos na seção de resultados são uma contribuição que pode permitir que professores os empreguem em aula. Trata-se de um contributo importante, porque muitas

vezes um professor pode não ter prontamente ideias acerca de quais são as incompreensões mais comuns dos estudantes. Assim, ter uma série de exemplos de concepções alternativas pode permitir que professores elaborem avaliações diagnósticas ou atividades de discussões. As discussões podem promover processos de mudança conceitual por metacognição, uma vez que as frases podem auxiliar a tomada de consciência de estudantes acerca de quais são as compreensões que cada um possui.

Por viabilizar a análise de falas dos participantes diretamente, foi possível discutir questões referentes à relação pensamento-linguagem tanto por meio do estudo das palavras utilizadas como das frases. Os termos metafóricos contribuem para visões equivocadas, porque são frequentemente utilizados de forma literal. Os adjetivos metafóricos "dominante" e "recessivo", originalmente utilizados por Mendel, foram entendidos de forma literal e foram desdobrados em substantivos e verbos. Ao utilizar o verbo "dominar", a frase passa a admitir um sujeito que se torna agente. Esse sujeito pode variar, sendo ou o alelo ou a característica. Estabelece-se então uma relação falsa de dominador-dominado entre dois alelos ou fenótipos. Consideramos, portanto, que a linguagem utilizada seja dificultadora da aprendizagem do conteúdo em questão.

Sinalizamos para alguns questionamentos referentes ao ensino: o uso da terminologia "alelo selvagem/mutante" ao invés de "alelo dominante/recessivo" no discurso didático poderia auxiliar no problema do desvio de nível do fenômeno da dominância do fenótipo para o alelo? O conteúdo de síntese proteica deveria preceder o ensino de genética de transmissão para subsidiar a compreensão acerca de como genótipos orientam a estruturação de fenótipos? A genética mendeliana com caráter histórico e conceitual precisa continuar sendo o primeiro conteúdo do ensino de genética?

Além de tais questionamentos, apontamos que as incompreensões constatadas nas explicações de professores em formação indicam a necessidade de articular os conteúdos de genética de transmissão e genética molecular durante a graduação de Ciências Biológicas. Novos estudos se fazem necessários para investigar como será possível reorganizar o estudo do material genético de modo a permitir a compreensão integrada dos processos de hereditariedade e dos processos de produção do fenótipo.

### **Agradecimentos**

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa concedida durante o doutorado em Educação para Ciência.

## Referências

- Allchin, D. (2000) Mending mendelism. *The American Biology Teacher*, 62 (9), 632-639. doi: [https://doi.org/10.1662/0002-7685\(2000\)062\[0632:MM\]2.0.CO;2](https://doi.org/10.1662/0002-7685(2000)062[0632:MM]2.0.CO;2)
- Bardin, L. (2011) *Análise de conteúdo*. São Paulo: Edições 70 LDA.
- Ceschim, B.; Ganiko-Dutra, M.; Caldeira, A. M. De A. (2020) Relação pensamento-linguagem e as distorções conceituais no Ensino de Biologia. *Ciencia & Educação*, 26 (1), 1-24. doi: <https://doi.org/10.1590/1516-731320200068>
- Ellis, T. H. N. et al. (2011) Mendel, 150 years on. *Trends in plant science*, 16(11), 590-596. doi: <https://doi.org/10.1016/j.tplants.2011.06.006>
- Flick, U. (2012). *Introdução à metodologia de pesquisa: um guia para iniciantes*. (M. Lopes, Trad.) Porto Alegre: Penso Editora.
- Garcia, B. H.; Pereira, T. C. (2016) O gene Stay-Green e a cor das sementes de Mendel. *Genética na Escola*. 11 (2), 386-391.
- Griffiths, A. J.; Wessler, S. R.; Lewontin, R. C.; Carrol, S. B.; (2011) *Introdução à genética* (P. A. Mota, Trad.) Rio de Janeiro: Guanabara Koogan.
- Heim, W. G. (1991) What is a recessive allele?. *The American Biology Teacher*, 53 (2), 94-97. doi: <https://doi.org/10.2307/4449229>
- Hull, D. (1975) *Filosofia da Ciência Biológica*. (E. de Almeida, Trad.) Rio de Janeiro: Zahar Editores.
- Klug, W. S.; Cummings, M. R.; Spencer, C. A.; Palladino, M. A. (2010) *Conceitos de Genética*. (9a ed., M. R. Borges-Osório, Trad.) Porto Alegre: Artmed Editora.
- Lewin, B. (2000) *Genes VII*. New York: Oxford University Press.
- Mahadeva, M. N.; Randerson, S. (1982) Mutation Mumbo Jumbo: Are mutations really so rare and harmful? Or are textbooks misleading?. *The Science Teacher*, 49 (3), 34-38.
- Marbach-Ad, Gili; Stavy, Ruth (2000) Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, v. 34, n. 4, p. 200-205. doi: <https://doi.org/10.1080/00219266.2000.9655718>
- Mendel, G. (2008) *Experiments in plant hybridisation*. [1909] New York: Cosimo Classics.
- Nachtomy, O.; Shavit, A.; Yakhini, Z. (2007) Gene expression and the concept of the phenotype. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 38, (1), 238-254. doi: <https://doi.org/10.1016/j.shpsc.2006.12.014>
- Nascimento-Schulze, C. M.; Camargo, E. V. (2000) Psicologia social, representações sociais e métodos. *Temas em Psicologia*, 8 (3), 287-299.
- Offner, S. (2011) Mendel's peas & the nature of the gene: genes code for proteins & proteins determine phenotype. *The american biology Teacher*, 73 (7), 382-387. doi: <https://doi.org/10.1525/abt.2011.73.7.3>

Pasternak, J. J. (2002) *Genética molecular humana: Mecanismos das doenças hereditárias*. São Paulo: Editora Manole.

Pavlova, I. V.; Kreher, S. A. (2013) Missing links in genes to traits: Toward teaching for an integrated framework of genetics. *The American Biology Teacher*, 75 (9), 641-649. doi: <https://doi.org/10.1525/abt.2013.75.9.4>

Pearson, J. T.; Hughes, W. J. (1988) Problems with the use of terminology in genetics education: 1, A literature review and classification scheme. *Journal of Biological Education*, 22 (3), 178-182. doi: <https://doi.org/10.1080/00219266.1988.9654979>

Pierce, B. (2011) *Genética: Um enfoque conceitual*. (3a ed., P. A. Motta Trad.) Rio de Janeiro: Guanabara Koogan.

Redfield, R. J. (2012) "Why do we have to learn this stuff?"—a new genetics for 21st century students. *PLoS Biol*, 10 (7), e1001356. doi: <https://doi.org/10.1371/journal.pbio.1001356>

Reece, J. B. et al. (2015) *Biologia de Campbell*. (10a ed., A. D. Villela et al., Trad), Porto Alegre: Artmed.

Reid, J. B.; Ross, J. J. (2011) Mendel's genes: Toward a full molecular characterization. *Genetics*, 189 (1), 3-10. doi: <https://doi.org/10.1534/genetics.111.132118>

Sampieri, R. H., Collado, C. F., & Lucio, P. B. (2013). *Metodologia de pesquisa*. Porto Alegre: Penso.

Schnur, R. E. et al. (1996) Type I oculocutaneous albinism associated with a full-length deletion of the tyrosinase gene. *Journal of investigative dermatology*, 106 (5), 1137-1140. doi: <https://doi.org/10.1111/1523-1747.ep12340185>

Seipelt-Thiemann, R. L. et al. (2012) Analogies for teaching mutant allele dominance concepts. *Creative Education*, 3 (6), 884-889.

Smith, M. U. (2014) It's not your grandmother's genetics anymore!(Part 2). *The american biology Teacher*, 76 (5), 306-310. doi: <https://doi.org/10.1525/abt.2014.76.5.3>

Smýkal, P. (2014) Pea (*Pisum sativum* L.) in biology prior and after Mendel's discovery. *Czech Journal of Genetics and Plant Breeding*, 50 (2), 52-64.

Thomas, H. et al. (1996) Gregor Mendel's green and yellow pea seeds. *Botanica Acta*, 109 (1), 3-4. doi: <https://doi.org/10.1111/j.1438-8677.1996.tb00862.x>

Thörne, K.; Gericke, N. M.; Hagberg, M. (2013) Linguistic challenges in Mendelian genetics: Teachers' talk in action. *Science Education*, 97 (5), 695-722. doi: <https://doi.org/10.1002/sce.21075>

Watson, J. et al. (2015) *Biologia Molecular do Gene*, (7a ed.) Porto Alegre: Artmed.