

遗传学教学中在细胞与分子水平上理解等位基因的显性与隐性

邢万金, 莫日根

内蒙古大学生命科学学院, 呼和浩特 010021

摘要: 在孟德尔遗传学中显性与隐性描述一对等位基因在杂合体时的功能关系, 把在表型上看出效果的等位基因称为显性等位基因, 看不出效果的称为隐性等位基因, 并由此提出分离定律和自由组合定律, 开创了遗传学这一学科。这样的描述最初是逻辑推理的需要, 但对于研究生命结构与功能关系的实验性科学而言, 必须在细胞与大分子实体上找到显性与隐性的生物学基础。在遗传学教学中, 如何用现代分子遗传学知识诠释经典的显性和隐性概念是教师经常面临的释疑问题。笔者认为要理解等位基因显性和隐性的实质, 必须了解等位基因的差异及其产物 RNA 或蛋白质在细胞中的具体作用。不同等位基因的蛋白质或者 RNA 产物在细胞内的不同时间、不同地点所起的作用不同, 赋予了在细胞、组织或器官水平上能够区分观测到的表型差异, 即显性或者隐性。文章根据基因结构的变异、基因调控的差异、基因产物的类型与作用等在细胞与分子水平上分别举例探讨了等位基因显性与隐性的分子实质及其变化, 以期在遗传学教学过程中使学生对基因的变异和功能有更全面、更具体的理解。

关键词: 遗传学; 教学; 等位基因; 显性; 隐性

Understanding the cellular and molecular mechanisms of dominant and recessive inheritance in genetics course

Wanjin Xing, Morigen

School of Life Sciences, Inner Mongolia University, Hohhot 010021, China

Abstract: In Mendellian genetics, the dominance and recessiveness are used to describe the functional relationship between two alleles of one gene in a heterozygote. The allele which constitutes a phenotypical character over the other is named dominant and the one functionally masked is called recessive. The definitions thereby led to the creation of Mendel's laws on segregation and independent assortment and subsequent classic genetics. The discrimination of dominance and recessiveness originally is a requirement for Mendel's logical reasoning, but now it should be ex-

收稿日期: 2014-09-02; 修回日期: 2014-10-01

基金项目: 内蒙古自治区生物化学系列课程教学团队建设项目资助。

作者简介: 邢万金, 博士, 教授, 研究方向: 分子遗传学和基因工程。Tel: 0471-4992944; E-mail: xwanjin@imu.edu.cn

莫日根, 博士, 教授, 研究方向: DNA 复制调控与细胞周期。E-mail: morigenm@life.imu.edu.cn

邢万金和莫日根并列一作者。

致谢: 非常感谢中山大学的贺竹梅教授和内蒙古师范大学侯占铭教授为本文提出了许多修改建议。

DOI: 10.16288/j.ycz.2015.01.014

网络出版时间: 2014-10-15 8:05:18

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20141015.0805.001.html>

plained by cellular and molecular principles in the modern genetics. To answer the question raised by students of how the dominance and recessiveness are controlled, we reviewed the recent articles and tried to summarize the cellular and molecular basis of dominant and recessive inheritance. Clearly, understanding the essences of dominant and recessive inheritance requires us to know the dissimilarity of the alleles and their products (RNA and/or proteins), and the way of their function in cells. The alleles spatio-temporally play different roles on offering cells, tissues or organs with discernible phenotypes, namely dominant or recessive. Here, we discuss the changes of allele dominance and recessiveness at the cellular and molecular levels based on the variation of gene structure, gene regulation, function and types of gene products, in order to make students understand gene mutation and function more comprehensively and concretely.

Keywords: genetics; teaching; alleles; dominance; recessiveness

显性(Dominance)与隐性(Recessiveness)是孟德尔遗传学的基础,也是遗传学中最抽象难懂的概念之一。孟德尔遗传学用显性和隐性描述一对等位基因(Allele)在一起时(杂合体)的功能关系,把在表型上看出效果的那个等位基因称为显性等位基因,看不出效果的那个称为隐性等位基因。这样的抽象描述对于尚不知基因为何物的早期遗传学工作者而言其实是一种武断的规定。但这个武断的定义却是孟德尔遗传学的关键,因为只有相信这个概念才能解释一对相对性状杂交的 3:1 分离或两对相对性状杂交的 9:3:3:1 自由组合结果。但显性与隐性并非某等位基因固有的属性,例如在 $A1$ 、 $A2$ 、 $A3$ 3 个复等位基因(Multiple alleles)中, $A1$ 与 $A2$ 在一起时显性,但如果与 $A3$ 在一起则可能是不完全显性(Incomplete dominance)、共显性(Codominance)甚至隐性。那么,等位基因的区别是什么?为什么等位基因会有显性与隐性之分?显性与隐性的生物学实质是什么?为什么有的等位基因有时显性、有时隐性或者介于显性与隐性之间呢?初学遗传学的学生们始终被这些问题困扰着,甚至在学完遗传学后仍然说不清楚。

孟德尔遗传学的显性和隐性是可观察到或者可测定出的宏观性状。分子遗传学告诉我们,基因本身是核酸序列,某等位基因的显性与隐性应该与该等位基因的核酸序列在特定的细胞环境中是否发挥了作用有关系。基因主要通过其产物 RNA 或蛋白质在细胞中发挥作用。因此要理解显性和隐性的实质,必须先理解蛋白质或 RNA 在细胞中如何起作用。不同的等位基因的蛋白质或者 RNA 产物在细胞内不同时间、不同地点所起的作用不同,赋予了其在细胞、组织或器官水平上能够区分观测到的表型差异,

即显性或者隐性。有学者也曾撰文初步探讨在分子水平上理解基因的显性和隐性问题^[1],本文根据基因结构的变异、基因调控的差异、基因产物的类型与作用等在细胞与分子水平上分别举例全面探讨了等位基因显性与隐性的实质,以期在遗传学教学中使学生深刻理解等位基因显性与隐性的分子生物学基础。

1 编码蛋白质的基因的显性与隐性

1.1 编码酶的基因

编码酶的基因的功能取决于所表达出来的酶蛋白是否具有催化活性。由于基因结构的变异,编码酶的隐性等位基因可能产生一个酶活性发生了改变或者完全丧失的酶蛋白、甚至根本不能表达出蛋白。如果用等位基因 A 表示能够产生有酶活性的酶蛋白等位基因, a 表示不能产生酶蛋白或者产生的蛋白已经丧失了酶活性的等位基因,那么杂合体 Aa 与纯合体 AA 表达的酶蛋白产物中均有“有活性的”的酶,能都表现出显性性状,而且往往是完全显性;而 aa 纯合体则没有蛋白产物或者该基因产物都是丧失了酶活性的蛋白,不能够催化相应的代谢反应,表现出隐性性状。这种现象也被称为单倍性充足(Haplosufficiency),即只要一个功能性基因就足以维持细胞所需要的正常功能状态。例如 Tay-Sachs 病是一种罕见的进行性神经退化常染色体隐性遗传病,患者从 3~6 个月就开始发病,通常在 4 岁前就死亡。人类 15q23-24 处的 $HEXA$ 基因编码一种溶酶体酶称为 β -N-乙酰己糖胺酶 A(β -N-acetylhexosaminidase A),能够清除神经细胞中多余的细胞膜组份 GM2 神经节苷脂(Ganglioside),而 Tay-Sachs 患者的两个

HEXA 等位基因都有某种突变导致所编码的 β 己糖胺酶 A 完全没有活性^[2], 不能除去神经细胞中多余的 GM2 神经节苷脂, 导致 GM2 神经节苷脂在神经细胞中迅速积累, 引起细胞死亡, 表现 Tay-Sachs 病。这类显隐性的关系对于编码代谢所需要的酶的等位基因而言非常普遍。遗传学教学经常使用的案例有苯丙酮尿症、尿黑酸尿症等。之所以表现完全显性, 即只需一个有功能的等位基因就能完全保证细胞所需的正常性状, 是因为在细胞水平上, 这类酶必需但并不需要大量; 在分子水平上这类酶也不是被当作原料消耗掉的, 而是作为工具能够反复发挥作用, 所以即使杂合体产生一半量的有功能酶就足够维持表型。

血型(Blood type)是基于红细胞表面的抗原对血液进行的划分。抗原有蛋白质、碳水化合物、糖蛋白和脂蛋白等。这些抗原是细胞内合成的物质, 受基因的控制。国际输血学会(International society of blood transfusion, ISBT)目前认定人类有 33 种血型系统 (<http://www.isbtweb.org/working-parties/red-cell-immunogenetics-and-blood-group-terminology/blood-group-terminology/>), 其中与临床关系最密切的是 ABO 血型和 Rh 血型系统。ABO 血型系统由红细胞表面的 3 种糖脂组成, 分别称为 A 抗原、B 抗原和 O 抗原(H 物质), 它们是由同族凝集原(Isoagglutinin)基因的 3 个等位基因控制, 分别为 I^A 、 I^B 和 i 。 I^A (AF134412.1, NM_02046)编码 354 aa 的 N 乙酰半乳糖胺转移酶(GalNAc transferase)(<http://www.uniprot.org/uniprot/P16442>), 催化在 H 物质的半乳糖环末端加上一个 N 乙酰半乳糖胺(N-acetylgalactosamine)形成 A 抗原; I^B 是 I^A 编码区的 C⁵²⁶→G、G⁷⁰³→A、C⁷⁹⁶→A 和 G⁸⁰³→A 4 个碱基变异导致 R¹⁷⁶→G、G²³⁵→S、L²⁶⁶→M 及 G²⁶⁸→A 4 个氨基酸改变^[3], 编码半乳糖基转移酶(Gal transferase), 催化在 H 物质的半乳糖环末端加上一个半乳糖(Galactose)形成 B 抗原; i 是由于 I^A 编码区的 G²⁶¹ 缺失, 引起第 87aa 后移码突变, 导致其蛋白产物变短, 丧失酶活性。由于 I^A 和 I^B 两个等位基因都能转录翻译出有功能的酶, 并分别把 H 物质转化为 A 抗原和 B 抗原, 所以基因型 $I^A I^B$ 的个体的红细胞表面有 AB 两种抗原, 表现为 AB 血型, $I^A I^B$ 表现为典型的共显性关系。

1.2 编码结构蛋白的基因

酶发挥功能主要靠“质”而不靠“量”, 但结构性蛋白作为构成细胞和组织的原料, 则需要足够的量。它们对表型的贡献既取决于其正确的结构, 也取决于其数量。

胶原蛋白是动物体含量最丰富的蛋白质, 约占身体总蛋白的 25%~35%, 主要存在于结缔组织中, 如血管、软骨、皮肤、肌腱、韧带等。编码胶原蛋白的基因家族有 40 多个基因^[4], 其中 COL1A1 和 COL1A2 基因的编码产物互相结合形成的 I 型胶原蛋白是胶原蛋白中最丰富的形式。有 300 多种结缔组织疾病是由于这两个基因中的很多突变所致^[5], 仅成骨不全症(Osteogenesis imperfecta)一种 I 型胶原蛋白病就与 800 多个点突变有关^[6]。胶原蛋白是动物的必需蛋白, 野生型胶原蛋白基因纯合体(AA)表型正常, COL1A1 或 COL1A2 的突变型纯合体(aa)不能产生有功能的胶原蛋白而无法存活, 而杂合体(Aa)虽有一个正常基因, 但仍表现各种程度的胶原蛋白病, 这种现象称为单倍性不足(Haploinsufficiency), 往往表现为不完全显性。这个例子说明结构蛋白基因维持正常表型需要足够的野生型等位基因产物, 但杂合体只有一个野生型等位基因, 其产物的量不足以维持细胞正常表型, 就会表现某种程度的缺陷。只要一个等位基因变异就表现出缺陷性状, 这种性状即为显性性状, 所以变异的等位基因反而成为显性等位基因。

1.3 编码功能性蛋白的基因

血红蛋白是除了冰鱼科(Channichthyidae)^[7]外所有脊椎动物运送氧气的蛋白。成年人的血红蛋白是 α 和 β 各两条链构成的异源四聚体($\alpha_2\beta_2$), 其分子功能是在肺部高氧环境中结合氧分子, 随血液流动把氧分子送到机体各处低氧的环境中释放。人血红蛋白分别由 α 和 β 两个基因家族编码, 血红蛋白基因的表型可以用“贫血”(缺氧)与否来描述。要给机体提供足够的氧气, 既需要足够的血红蛋白(量), 又要求 α 和 β 单体的结构正常(质)才能结合足够多的氧分子。但另一方面, 血红蛋白能够结合多少氧气也与环境中的氧气浓度有关。这就使血红蛋白等位基因的显性与隐性关系变得复杂了。

镰刀形细胞贫血症(Sickle cell anemia)是人类第

一个在分子水平了解透彻的疾病。正常 β 血红蛋白链(HbA)的第 6 位氨基酸是谷氨酸,密码子突变后谷氨酸变成了缬氨酸就成为引起镰刀形细胞贫血症的 β 血红蛋白(HbS)。这个氨基酸位于 β 血红蛋白分子的空间构象的表面,野生型的谷氨酸是极性氨基酸(有侧链羧基),亲水,使血红蛋白易溶于水,而突变型的缬氨酸却是非极性氨基酸(侧链是甲基),疏水。HbS 比 HbA 的氧气亲和能力差,更容易释放所携带的氧^[8],当血红蛋白把携带的氧分子释放掉以后会发生构象变化,暴露出由 85 位苯丙氨酸和 88 位亮氨酸形成的一个疏水性口袋,而 HbS 的第 6 位缬氨酸正好在外侧,就会插入邻近的失去氧气的 HbS 血红蛋白的疏水性口袋^[9],于是很多 HbS 蛋白互相聚集成不溶性沉淀,使红细胞变形失去弹性,不能通过毛细血管,引起贫血。

HbS 纯合体(HbS/HbS)往往由于严重的贫血而死亡,但杂合体(HbA/HbS)一般不表现贫血,说明 HbA 对于 HbS 等位基因是显性,即单独一个正常 HbA 等位基因的产物 HbA 能够提供细胞所需要的氧气,表现为单倍性充足。但当环境中氧气稀薄(如高原)或者机体短时间需要大量氧气供应(如高强度劳动或体育竞技)的时候,杂合体(HbA/HbS)也会表现贫血,因为这时单独一个正常 HbA 等位基因的产物 HbA 的量已不能提供机体所需要的足量氧气, HbA 结合氧气的机会也变少,暴露出的疏水性口袋也会与相邻的 HbS 结合,导致整个血红蛋白复合体的疏水性口袋都与相邻的 HbS 结合成团,表现出与 HbS 纯合体(HbS/HbS)相同的镰刀形细胞贫血症,严重时甚至死亡。此时 HbA 等位基因表现出单倍性不足,而 HbS 基因则变成了显性等位基因。

判断某等位基因的显性与隐性主要依据与该等位基因所关联的性状。某个基因可能与多种性状有关,讨论所关联的不同性状的时候,等位基因的显隐性关系也不同。对于镰刀形细胞贫血症性状, HbS 等位基因一般为隐性等位基因。但杂合体(HbA/HbS)对疟疾的抗性却比野生型纯合体(HbA/HbA)强,因此在疟疾泛滥的地区(一般位于低海拔), HbA/HbS 杂合体不但不表现贫血病性状,反而因为抗疟疾而比野生型 HbA 纯合体的适合度更高^[10],因而 HbS 等位基因对于抗疟疾性状而言是显性,而 HbA 则为隐性。不过, HbS 纯合体(HbS/HbS)虽然感染疟疾的发病率

确实更低,但由于这种基因型即使在低海拔地区也会贫血,再加上疟疾本身也引起贫血,所以还是由于严重的贫血而比杂合体死亡率更高^[11]。前述的 Tay-Sachs 病也是一个类似的例子, Tay-Sachs 杂合体对结核菌有抗性^[12],对于抗结核病这个性状而言,突变的 *HEXA* 等位基因反而变成了显性等位基因。

细胞中还有一类非常重要的功能性蛋白称为转录因子(Transcription factors, TFs)。对于编码转录因子的基因而言,其基因产物虽然往往不是酶,但其功能在某些方面类似于酶的作用,即在发挥作用的过程中本身并不被消耗,所以其显性和隐性效果主要由转录因子的“质”决定,并不要求基因产物的数量。杂合体带有一个有功能的等位基因,所编码的转录因子能够与靶基因的启动子有效结合并激活靶基因表达,所以对突变的等位基因而言是简单的显性。

但有些转录因子的功能形式是相互结合的复合体,情况就不同了。如转录因子 p53 在 DNA 损伤的诱导下会迅速增加表达,激活下游的基因表达,使发生了 DNA 损伤的细胞停止分裂,给修复系统留出时间修复 DNA 或者让带有 DNA 损伤的细胞凋亡,因而被称为基因组的卫士(Guardian of the genome)^[13]。如果 p53 基因缺失,细胞不能产生 p53 蛋白,或者 p53 基因突变,产生的 p53 蛋白没有功能的话,细胞就会由于其内的 DNA 损伤得不到有效修复而容易发生癌变。

野生型的 p53 蛋白只有同源四聚体的形式才有功能^[14]。野生型 p53 基因纯合体(+/+)的所有 p53 蛋白单体都是正常的,全部能够形成同源四聚体,因此有转录因子的活性。但 *Tp53* 基因杂合体(+/*p53*)带有一个突变型 p53 等位基因,表达约一半数量的异常 p53 蛋白单体,导致形成的四聚体里往往会混入一个或多个异常 p53 蛋白单体,难以形成有功能的 p53 四聚体,或形成的杂合 p53 四聚体没有功能,使 p53 失去转录因子的作用^[15],表现为 dominant negative(显性抑制),因而 *Tp53* 基因杂合体(+/*p53*)细胞容易癌变,对于细胞癌变性状而言,突变的 p53 等位基因反而成为显性等位基因。

2 非编码基因的显性与隐性

细胞内的基因除了编码蛋白质的基因(称为编码基因)外,还有一些基因并不编码蛋白质,而是通过

转录出的 RNA 产物发挥功能, 称为非编码基因 (Noncoding genes), 如 rRNA、tRNA 基因以及近年来发现的大量非编码 RNA 基因, 如 MicroRNA、lncRNA 等 RNA 基因。rRNA 和 tRNA 基因是多拷贝的, 依靠产物的量发挥作用。人类 rRNA 有 28S、5.8S、5S 和 18S 4 种, 其中 28S、5.8S、18S 是从一个 45S 的前体中剪切出来的。人类共有 5 个 45S 基因簇, 每个簇中包含 20~30 个重复单元; 5S rRNA 簇中有 200~300 个真基因以及一些假基因。由于 rRNA 都是构成核糖体的组份, 因此单独某一份 rRNA 等位基因突变后即使其转录的 rRNA 不能组装到核糖体, 但多数其他正常的 rRNA 仍能保证细胞内足量核糖体的组装, 所以不会影响到细胞内蛋白质的合成, 因此突变的 rRNA 基因是隐性的。Genomic tRNA Database(<http://gtrnadb.ucsc.edu/Hsapi19/>)确认人类有 506 个编码 20 种标准氨基酸的 tRNA 基因。可以看出对于 64 个密码子, 平均解读每个密码子的 tRNA 有 10 多个基因, 因此情况与 rRNA 类似, 突变的某个 tRNA 等位基因一般是隐性的。

真核生物还有一些非编码 RNA(ncRNA), 如 snoRNAs、microRNAs、siRNAs、snRNAs、exRNAs、piRNAs 以及长 ncRNAs(如 *Xist* 和 *HOTAIR*)等。近年来发现小 RNA 除了在内含子剪切和 DNA 复制^[16]过程中起作用外, 还能调节众多的编码蛋白质基因的表达^[17]。miRNA 通过与 mRNA 序列配对使 mRNA 降解或抑制其翻译, 对编码基因进行负调控。每种 miRNA 可能调节多种靶基因, 几十种 miRNA 就可能调节上千种蛋白质基因的表达^[18], 因此它们的突变可能影响多种表型。讨论这些 miRNA 基因的显性与隐性需要根据与其相关的某些明显的表型来确定。但鉴于这些 miRNA 所调节的靶基因的多样性, 所引起的表型也复杂, 其显性或隐性只能就某一种特定的表型而言。

3 基因表达改变引起的显性与隐性变化

一般而言等位基因被认为是来源于某一个基因的不同功能的突变形式。其显性与隐性是由于等位基因之间在 DNA 序列上存在差异, 比如点突变和缺失突变等, 并造成所表达的基因产物在功能上出现区别。如果基因序列并不改变, 而是基因表达模式

发生了变化, 也会影响基因的功能和表型, 这种情况并非传统意义上的等位基因, 被称为表观等位基因(Epialleles)^[19], 也能区分显性与隐性。在表达水平上影响基因功能的方式有多种, 如启动子改变、选择性剪接(Alternative splicing)、表观遗传变异(Epigenetic variation)、小 RNA 调控等。

3.1 启动子改变与等位基因的显性和隐性

高等真核生物体细胞所含的基因组相同, 但不同组织细胞的基因表达模式不同。有些基因参与细胞基本生命活动, 在各种类型的细胞中均表达, 被称为管家基因(House-keeping genes)。但更多的基因是在特定的组织中或某组织发育的特定时段表达, 因而决定特定的表型, 被称为组织特异性基因(Tissue-specific genes)。表达模式的区别取决于该基因的启动子, 如果启动子发生了改变, 基因就可能不表达或者在其他组织和其他时间也表达, 影响表型。因此讨论某等位基因的显性与隐性还应该考虑该基因的表达模式改变而引起的表型变化。

小鼠 agouti 毛色(胡椒面色)表型主要由毛囊黑素细胞合成的色素所决定, 小鼠 α 毛囊黑素细胞刺激激素(α -melanocyte stimulating hormone, α -MSH)与 *extension* 基因编码的黑皮质素受体 1(Melanocortin 1 receptor, Mclr)结合, 促使毛囊黑素细胞合成黑色素(Eumelanin), 使毛成为黑色。控制小鼠 agouti 毛色的 agouti 信号蛋白(Agouti signal protein, ASP)基因 *ASP* 是个显性基因, 编码 ASP 信号蛋白。ASP 蛋白在毛发生长周期的第 4~6 天集中产生, 竞争性地与 Mclr 结合, 促使毛囊黑素细胞转而合成黄色的棕黑素(Pheomelanin), 第 6 天以后 *ASP* 基因停止表达, 毛囊黑素细胞恢复产生黑色素, 导致每根毛的末梢附近有一小段棕黄色的环带, 整体皮毛效果呈现 agouti(胡椒面色)表型^[20]。*ASP* 基因有很多等位基因, 有的等位基因是把 *ASP* 基因置于其他组成型表达启动子下, 如 *A^v* 等位基因。*ASP* 基因位点有 3 个基因, 从 5'至 3'依次为 *Raly*、*Eif2s2* 和 *ASP* 基因。*Raly* 基因的编码区、*Eif2s2* 基因的全部以及 *ASP* 基因的 5'调控区共约近 170 kb 缺失使 *ASP* 基因的编码区置于 *Raly* 基因的 5'调控区之下就成为 *A^v* 等位基因。*A^v* 等位基因杂合体中, 由于 *Raly* 基因的启动子导致 ASP 蛋白在各种组织中持续表达, 杂合体的毛色全长达

成为黄色^[21], 黄色毛对 agouti 毛色呈显性。尽管 agouti 蛋白的编码区没有变异, 但启动子被替换, 表达方式发生了改变, 本来显性的 ASP 等位基因变成了隐性。不仅如此, 组成型启动子也驱动 ASP 蛋白在身体的其他组织里持续表达, 结果使老鼠肥胖、容易发生肿瘤等, 引发了新的显性性状^[22]。

3.2 选择性剪接与等位基因的显性和隐性

选择性剪接(Alternative splicing)是真核基因转录出 mRNA 前体(pre-mRNA)后, 剪掉 pre-mRNA 中的内含子, 拼接外显子的时候选择性地保留或者剔除某些内含子或外显子, 导致同一个基因的不同外显子组合编码不同的蛋白质, 从而发挥不同的作用。选择性剪接在真核生物中是一个普遍现象, 它使有限数目的基因能编码出更多种类的蛋白质^[23]。具有多个内含子的人类基因中约 95% 的基因具有选择性剪接现象^[24]。最近也发现一些等位基因特有的选择性剪接(Allele-specific alternative splicing, ASAS)现象^[25], 提示某些等位基因可能有自己特定的选择性剪接模式, 从而表达不同的蛋白质, 有可能引起表型的改变甚至该等位基因的显性与隐性的转换。

Wnt 信号途径是哺乳动物胚胎期脊柱中轴形成的重要调控因子, 该信号途径中有一个 Axin 蛋白, 由具有 8 个外显子 *Axin*(axis inhibition)基因编码。小鼠有一种扭结尾巴(Kinked tail)变异是由于 17 号染色体上 *Axin* 基因的突变形式之一 *Axin^{Fu}*(*axin-fused*)引起的尾椎骨椎体融合而成^[26]。*Axin^{Fu}* 是个显性等位基因, 是在野生型 *Axin* 基因的第 6 内含子中插入了一个逆转座子 IAP(Intracisternal-A particle)所致, IAP 的插入会引起 *Axin^{Fu}* pre-mRNA 选择性剪接^[27], 丢失了第 7、8 外显子, 翻译出剩下的 1~596 aa 的短 *Axin^{Fu-NT}* 蛋白, *Axin^{Fu-NT}* 蛋白会与 832 aa 的野生型 *Axin* 蛋白形成异源二聚体, 使野生型 *Axin* 蛋白不能再激活 JNK, 表现出显性抑制效应^[28], 导致 *Axin^{Fu}/+*小鼠的 kinked 尾巴变异。有趣的是, IAP 插入并非导致全部的 *Axin^{Fu}* pre-mRNA 发生选择性剪接, 即使 *Axin^{Fu}/Axin^{Fu}* 纯合体也产生部分正常剪接的 *Axin* mRNA, 所以有些 *Axin^{Fu}/Axin^{Fu}* 鼠胎死腹中, 有些能带着 kinked 尾巴活到成年, 并伴有耳聋和行走行为异常。此外, 由于 IAP 部位的 DNA 甲基化等表观遗传修饰(Epigenetic modification), *Axin^{Fu}/+*小

鼠的 kinked 尾巴变异在群体中呈现不完全外显(Penetrance)和表现度(Expressivity)^[29,30]。

遗传学教科书中的控制果蝇性别的 *sxl*(*Sex lethal*)基因是一个选择性剪接控制不同性别的例子。当果蝇的 X 染色体条数与常染色体套数比值(X:A, 性指数)为 1 时, *sxl* 基因的 8 个外显子中的第三外显子被剪掉, 剩下的 7 个外显子拼接成为一个较大的多肽, 引发下游的 *tra*(*transformer*)基因的 4 个外显子也发生选择性剪接, 剪掉带有终止密码子的第二外显子, 拼接后的翻译产物继而引发另一个基因 *dsx*(*doublesex*)的 6 个外显子也发生选择性剪接, 剪掉第五和第六两个外显子, 拼接后的 *dsx* 蛋白诱导果蝇的性腺发育成卵巢, 成为雌性果蝇。而当果蝇的性比为 0.5 的时候, *sxl* 基因的 8 个外显子都保留, 但第三外显子内部会遇到一个终止密码子, 导致翻译提前终止, 反而表达出一个较小的蛋白质, 而这时下游 *tra* 基因的 4 个外显子也都保留, 但其第二外显子内部也会遇到终止密码, 导致翻译不出有功能的 *tra* 蛋白, 于是, 下游的 *dsx* 基因的 6 个外显子发生另一种形式的选择性剪接, 即剪掉第四外显子, 剩下的 5 个外显子拼接翻译出较大的 *dsx* 蛋白, 诱导果蝇的性腺发育成精巢, 成为雄性果蝇。此例中果蝇 *sxl*、*tra* 和 *dsx* 等基因决定的性状是性别, 尽管它们并非像前述的 *Axin^{Fu}* 那样在杂合状态下具有等位基因特异性选择性剪接差异, 但同一个基因的不同剪接产物可引发不同的性状已是不争的事实。

雌性玉带凤蝶(*Papilio polytes*)的翅膀斑点拟态红珠凤蝶(*Pachliopta aristolochiae*), 而雄性不拟态。最近发现, 玉带凤蝶的 *dsx* 基因也是通过选择性剪接决定性别, 但这个基因有 3 种不同的选择性剪接方式, 除了一种剪接产物在身体里表达影响雌雄性别外, 还有两种剪接产物在雌性的翅膀中表达, 决定了雌性玉带凤蝶翅斑点的拟态效应^[31], 因此这个基因的某种选择性剪接形式成为决定雌性拟态的“显性等位基因”。

最近发现人类的 *CACNA1A* 基因的 mRNA 除了翻译出钙通道蛋白的一个亚基 $\alpha 1A$ 外, 还能借助其内部核糖体进入位点(IRES)翻译出另一种结构和功能与之完全不同的多肽 $\alpha 1ACT$ ^[32]。正常的 $\alpha 1ACT$ 是一个转录因子, 激活 Purkinje 细胞的若干个基因表达, 促进神经轴突生长。变异的 $\alpha 1ACT$ 包含了更

多的多聚谷氨酰胺簇(PolyQ tract), 引发脊髓小脑性共济失调 6 型(Spinocerebellar ataxia type 6, SCA6)病症。这个基因的两种蛋白产物并非其 mRNA 的选择性剪接所致, 是一个真正的双功能基因(Bifunctional genes), 这样的基因给我们区分等位基因并讨论其显隐性更增加了复杂性。

3.3 表观遗传修饰与等位基因的显性和隐性

表观遗传变异(Epigenetic variation)是指基因的 DNA 序列本身不变, 但基因功能发生了可遗传的变化。引起表观遗传变异的主要原因是 DNA 甲基化和组蛋白修饰。容易发生这种表观遗传变异的等位基因称为亚稳定表观等位基因(Metastable epialleles), 简称表观等位基因(Epiallele), 能够稳定地以多种表观遗传状态存在^[33], 引起基因功能改变, 导致不同的表型^[34]。

表观遗传修饰能有效地改变基因的表达, 但并非永久性改变, 而是一种动态的变异, 其 DNA 甲基化或组蛋白去乙酰化等变化会受到调控, 甚至受到营养环境的影响。前述小鼠的 *agouti* 基因有一种变异是在 *ASP* 基因编码区上游插入了一个逆转座子 IAP, 称为 *A^{iapy}* 等位基因, 该逆转座子的长末端重复(LTR)具有启动子效应, 驱动 *agouti* 基因的编码区持续表达, 导致 *A^{iapy}* 杂合体毛全长黄色, 因而 *A^{iapy}* 对野生型 *ASP* 等位基因(*A*)显性。但 IAP 的 LTR 上的 CpG 会受到甲基化, LTR 处的组蛋白 H4K20 也甲基化^[35], 从而会抑制 LTR 启动子的效率。如果胚胎期获得富含甲基的营养成分, 会导致出生后有些 *A^{iapy}A* 杂合体的 *A^{iapy}* 等位基因受到甲基化抑制, 预期的黄毛色(显性)变成了 *agouti* 色(隐性), 称为假 *agouti*^[36], 此种个体中 *A* 对 *A^{iapy}* 成为显性等位基因。说明表观遗传修饰会改变基因的表达从而影响等位基因的显性与隐性。

基因组印记(Genomic imprinting)是一种特殊的表观遗传变异, 真核生物有些基因的表达与否取决于它们来自父本还是来自母本。子代的同一个等位基因如果来自父本可能有功能, 而来自母本则沉默, 反之亦然, 这使得区分显性与隐性更为复杂。

1983 年俄克拉荷马州的牧羊人在自己的羊群中筛选肌肉量增加的羊, 发现一只 Dorset 公羊拥有出众的大臀部, 臀部及后腿中肌肉量显著增多, 而脂

肪减少, 取名为“Solid Gold”。“Solid Gold”与野生型母羊交配后能产生 50% 大臀羊, 雌雄都有, 显示这是个常染色体显性单基因遗传, 并把这些大臀羊称为 callipyge(拉丁文“美臀”的意思), 该基因座位后来被命名为 *CLPG*^[37]。但 1980 年代中后期发现用 callipyge 母羊与野生型公羊交配却不产生 callipyge 后代, 只有杂合体 callipyge 公羊与野生型母羊交配的后代中才有 callipyge, 说明 *CLPG* 基因在母羊中存在基因组印记。但选用 *CLPG* 等位基因都是来自杂合体 callipyge 公羊(*CLPG^{Pat}/+*)的杂合体 callipyge 羊互相交配(上标 Pat 表示该等位基因来自父本), 后代中并不是预期的 50% callipyge, 而是 75% 野生型和 25% 的 callipyge, 基因分析显示, 后代中只有杂合体 *CLPG^{Pat}/+* 表现 callipyge, 而 *CLPG^{Pat}/CLPG^{Mat}* 纯合体却表现为野生型(上标 Mat 表示该等位基因来自母本), 与野生型纯合体 *+/+* 的表现型一样! 即来自父本与来自母本的同一个等位基因的显性与隐性不同, *CLPG^{Pat}* 对野生型等位基因是显性, *CLPG^{Mat}* 对野生型等位基因是隐性, 但当它们在一起时(*CLPG^{Pat}/CLPG^{Mat}*), 显性和隐性会转换, *CLPG^{Mat}* 对 *CLPG^{Pat}* 等位基因成为显性, 这种现象被称为 polar overdominance(极性超显性)^[38]。

经过鉴定, 上述 *CLPG* 基因座位是一个约 1 Mb 长的保守的印记结构域, 其中至少有 4 个雄性优先表达的编码蛋白质的印记基因: *BEGAIN*、*DLK1*、*PEG11* 和 *DIO3*, 还有雌性优先表达的非编码 RNA(noncoding RNA, ncRNA) 基因: *GTL2*、*anti-PEG11*、*MEG8* 和 *MIRG* 以及 C/D snoRNAs 和 miRNAs。DNA 测序发现, *CLPG* 等位基因与野生型的区别实际上只是 *DLK1* 与 *GTL2* 基因之间的 DNA 序列中的有一个碱基转换突变(A→G)^[39,40]。由于这个突变并非发生在这些基因的内部, 也并不影响上述基因的印记模式, 推测该突变可能位于某种负调控元件中。科学家鉴定了这些基因在不同的 *CLPG* 基因型的骨骼肌细胞中的转录情况, 发现 *CLPG^{Pat}/+* 和 *CLPG^{Pat}/CLPG^{Mat}* 基因型的羊都是 *DLK1* 和 *PEG11* 的 mRNA 超表达, 而 *GTL2*、*antiPEG11*、*MEG8* 及 *MIRG* 的 RNA 水平与野生型纯合体 *+/+* 并无区别, 即在转录水平上并未发现与 callipyge 表现型有关联的特征, 但当鉴定蛋白质的时候, 发现 *DLK1* 蛋白在 *CLPG^{Pat}/+* callipyge 羊的骨骼肌中大

量存在,而在 $CLPG^{Pat}/CLPG^{Mat}$ 、 $+/+$ 和 $CLPG^{Mat}/+$ 基因型的羊骨骼肌中检测不到,说明可能是 *DLK1* 蛋白引起 callipyge 表现型^[41]。野生型 $CLPG(+)$ 可能是个沉默子,使这个座位的基因维持适度的表达,而突变后的 $CLPG^{Pat}$ 等位基因导致两侧的 *DLK1* 和 *PEG11* 基因转录增强,但 $CLPG^{Mat}$ 等位基因却导致 *antiPEG11* 和 *MEG8* 等 ncRNA 表达增加,这些 ncRNA 会抑制 *DLK1* 和 *PEG11* mRNA 的翻译,因此 $CLPG^{Pat}/CLPG^{Mat}$ 基因型中父源 $CLPG^{Pat}$ 等位位点上增加的 *DLK1* 和 *PEG11* mRNA 被母源 $CLPG^{Mat}$ 等位位点上增加的 ncRNA 抵消了,不能翻译出比野生型更多的 *DLK1* 蛋白,无法表现 callipyge 表现型; $CLPG^{Mat}/+$ 基因型只是 ncRNA 增加,并未显著影响羊的臀部肌肉量;只有 $CLPG^{Pat}/+$ 基因型增加了 *DLK1* 和 *PEG11* 的 mRNA,且未受更多的 ncRNA 抑制,得以翻译出更多的 *DLK1* 蛋白,表现出 callipyge 表现型^[42]。为何一个单核苷酸多态性 (SNP) 变化就引起如此大的效应? 人们发现突变的 $CLPG$ 位点的 CpG 甲基化降低 (Hypomethylation) 改变了该处染色质的结构,促进两侧的基因转录^[43,44]。

4 哺乳动物性染色体上等位基因的显性与隐性

二倍体哺乳动物常染色体上的等位基因都有杂合基因型,能够区别显性与隐性。但性染色体上的基因就不易区分显性与隐性。人类的正常性染色体组成是女性 XX, 男性 XY。X 与 Y 并非同源染色体,彼此只有假常染色体区 (Pseudoautosomal regions, PARs) 同源,所涉及的基因不多,大部分区域不同源。所以男性的 X 和 Y 染色体上的绝大多数基因是单倍型的,不存在隐性状态。

女性的 XX 是一对同源染色体,其上会携带有等位基因杂合体,理论上具有显性和隐性之分。但实际上根据莱昂假说,女性体细胞中的两条 X 染色体只有一条有活性,另一条失活,所以对每一个细胞而言,其中的 X 连锁基因也都是单倍型的,不存在隐性状态。这样的情况对于 X 连锁基因为纯合体 ($X^A X^A$ 或 $X^a X^a$) 的女性而言与男性相同,由于每个细胞中的两条 X 上所带的等位基因相同,每个细胞虽然只留下一条有活性的 X 染色体,但所带的 X 连锁基因也相同 (X^A 或 X^a)。

然而,对于 X 连锁基因的杂合体 ($X^A X^a$) 女性而言,情况并非如此简单。每个细胞中两条 X 染色体中的一条在胚胎早期开始随机失活,而且一旦失活,这个细胞分裂后的所有后代细胞都会保持这条失活的 X 染色体继续为失活状态。这种失活的结果是女性的某些体细胞中携带显性等位基因的 $X(X^A)$ 有活性,而有的体细胞却是携带隐性等位基因的 $X(X^a)$ 有活性。实际上导致 $X^A X^a$ 女性是由保留活性 X^A 的细胞群与保留活性 X^a 的细胞群组成的嵌合体。该女性的表型将由 X^A 细胞群与 X^a 细胞群的数量、分布以及表型决定。由于大部分编码蛋白质的基因是组织特异性表达的基因,因此 A 和 a 能否在女性的整体表型上显示出来,取决于 X^A 细胞群与 X^a 细胞群在各种器官和组织中的分布。如红绿色盲是 X 连锁隐性遗传病,只在视网膜中表现性状。杂合体 ($X^A X^a$) 女性如果其视网膜组织细胞中的 X^A 失活,而留下 X^a ,则表现红绿色盲。此时杂合体个体反而表现了隐性等位基因的性状,隐性等位基因看起来成了显性。

上述红绿色盲表型只与视网膜细胞有关,其他体细胞组织中的红绿色盲基因并不影响视觉性状,所以不论活性 X^a 的细胞数目在整个体细胞中有多少,只要恰好在视网膜细胞中占多数就显示出红绿色盲。但对于那些在体细胞中普遍发挥作用的 X 连锁基因而言,情况有所不同。由于 X 染色体随机失活,而且发生在胚胎发育的早期,就有可能导致 $X^A X^a$ 女性体细胞中的留下活性 X^a 的细胞占了多数,X 连锁隐性基因在杂合体整体上表现出性状,隐性等位基因看起来成了显性。例如人类的 Lesch-Nyhan 综合征是 X 连锁基因编码的次黄嘌呤鸟嘌呤磷酸核糖转移酶 (Hypoxanthine guanine phosphoribosyltransferase, HPRT) 功能缺陷造成的 X 连锁隐性遗传病,有的 $X^+ X^{HPRT}$ 女性由于 X^+ 失活的细胞多,留下活性 X^{HPRT} 的细胞占了多数,表现出 Lesch-Nyhan 病症,甚至还发现过一对同卵双胞胎的 $X^+ X^{HPRT}$ 女性,其中一个表现出 Lesch-Nyhan 病症,而另一个正常^[45]。本例中 $X^+ X^{HPRT}$ 表现 Lesch-Nyhan 病症并非丧失功能的 HPRT 基因产物真的“战胜”了野生型等位基因编码的有功能的 HPRT 酶,而是由于多数体细胞内带有野生型等位基因的 X 染色体失活而缺乏有功能的 HPRT 酶,少数体细胞中仍然保留野生型等位基

因活性,但这些细胞并无可区分的相对性状,因此在整体上表现出 Lesch-Nyhan 病症状,即杂合体表现了隐性性状。类似的例子还有玳瑁猫,玳瑁猫躯体的毛色也是 X 连锁基因控制的,雌性杂合体猫的毛色斑点分布是由于 X 染色体随机失活所致,由于 X 连锁隐性等位基因和显性等位基因都控制可区分的毛色相对性状,在整体上看起来似乎是一对等位基因共显性形成的无规则色斑。

5 基因组和环境对显性和隐性的影响

等位基因构成的基因型通过显性或者隐性决定表型,环境通过筛选表型来筛选基因型。严格地说所有的表型都是由多种基因甚至是整个基因组的所有基因在细胞内相互作用并在组织器官上显示出来的性状,其实根本不存在“单基因性状”。我们所讨论的某等位基因控制的显性性状或隐性性状在绝大多数情况下是默认与该性状相关的其他基因(遗传背景)没有变化。但与之相关的其他基因甚至是外界环境的变化,也会影响我们所讨论的等位基因的显性或者隐性。

在遗传学教学中,上位效应(Epistasis)描述了某基因能否实现其所控制的性状依赖于基因组中其他基因的功能是否实现。例如核糖体蛋白和一些信号通路或代谢通路中的蛋白,都能正确表达才能协作实现最终表现型,如果其中一个基因丧失了功能,则其他基因即使有正常的产物也不能表现性状,都变成了隐性。还有一些基因之间的联系虽然不如代谢通路或者信号通路中的基因联系这么紧密,但某些基因的功能为其他基因发挥功能提供了平台,如秃顶基因与头发的颜色基因,女性性腺发育相关基因与乳腺癌、卵巢癌基因等。

环境也可能诱导或抑制基因的表达,从而影响基因的显隐性。如前述的小鼠 agouti 显性等位基因 A^{iapy} , $A^{iapy}A$ 杂合体小鼠如果在胚胎期间获得大量的富含甲基的营养物,其 LTR 启动子的 CpG 被甲基化,抑制了该显性等位基因的活性,毛色表现为隐性等位基因 A 的 agouti 性状(假 agouti)。

环境变化能够诱导激活或者关闭某些基因,甚至其表达方式发生永久性的改变。人类血红蛋白负责为机体提供氧气。血红蛋白基因是一个超家族,不同的家族成员在人生发育的不同阶段起作用。两

个 α 和两个 β 血红蛋白组成的四聚体($\alpha_2\beta_2$)是出生后成年人的血红蛋白功能形式, α 和 β 家族的另外几个成员组成的四聚体在胚胎发育过程中负责从母体胎盘中为胚胎和胎儿收集氧气,如 $\alpha_2\epsilon_2$ 是胚胎血红蛋白, $\alpha_2\gamma_2$ 是胎儿用的血红蛋白。为什么同一个表型需要不同的基因呢?这跟出生前后的环境有关系。胎盘绒毛里的氧气比母体动脉血管里的氧气浓度低,胎儿血红蛋白与氧气的亲和力必须比成人血红蛋白高,才能够从胎盘里有效地收集氧气。但出生后用肺呼吸,接触空气中高浓度的氧气,胎儿血红蛋白的高亲和力反而有害。因为机体活动,组织缺氧,需要血红蛋白提供氧气,但胎儿血红蛋白与氧亲和力高,不容易给缺氧组织快速释放氧气,不能适应在体外快速活动的需求。因此在不同的环境中,可能会关闭掉本来显性的基因,而开启原本不表达的基因。

6 结 语

显性和隐性是遗传学中最基础也是最抽象的概念之一。我们借助现代生物技术已能够看到 DNA,但看不到“基因”,也无法描述什么样的 DNA 是基因,因为“基因”是某一段 DNA 的功能属性。我们也看不到“显性”或“隐性”,所以我们无法给学生展示什么样子是显性,什么样子是隐性。但在遗传教学中我们确实常说某种性状是显性或隐性的,也说某等位基因是显性或隐性的,这些本来就抽象的概念纠结在一起更加难以理解。其实显性/隐性并非描述基因的结构变化,比如突变、缺失等,它与基因的突变有关系,但并不对应,它实际上是描述基因的功能状态。“功能”不能被看到,只能根据在某个过程后的某种结果的变化去判断,而且如何描述也取决于我们如何定义“功能”,尤其是在哪个水平上定义。

“结构”是具体的、静态的,但“功能”是抽象的、动态的,离开了“结构”也就不会有该结构的功能,但该“结构”存在并不一定总能表现出“功能”,或者表现出“相应的功能”。基因能不能表现出功能,受制于其他基因和整个细胞、整个有机体甚至外界环境的影响。孟德尔遗传学在有机体的层面上根据表型的变化宏观地定义了显性和隐性,有助于我们尽快地找到问题并提炼出宏观变化规律,

是一种非常成功的科学研究办法。但仅仅停留在这种粗糙的观察和逻辑推理的话, 就会发现很多情况让人费解甚至误解。分子遗传学告诉我们, 基因只是在分子和细胞水平上发挥作用, 而且并非所有的基因都能在个体水平上表现出其独特的性状, 观察到的性状也不仅仅是某一对等位基因的显性或隐性所致, 某一基因甚至有多种产物, 因而影响多种形状, 强调我们必须深入到细胞或分子水平上分析基因产物及其真正的功能才能了解基因显性与隐性的本质。如果我们把某等位基因在分子水平或细胞水平上“发挥了作用”定义为显性, 而相对的等位基因形式在某一条件下“没发挥出作用”或其“作用被掩盖”定义为隐性, 就能始终把握基因及其功能的实质, 全面理解表型与基因型的关系, 不仅能够分子遗传学教学中继续使用显性和隐性概念, 凸显遗传学教学的连贯性和由浅入深由表及里的遗传分析逻辑, 而且能够用这个概念指导学生更好地理解基因的功能, 有利于学生把基因的结构、变异、转录、加工、调控、翻译等分子遗传学的核心内容与之之前学过的等位基因及其显隐性联系起来, 搭建微观与宏观之间的联系桥梁, 理解基因在生命活动中的必要性和与其他基因相互作用的复杂性, 这才是遗传学教学的最终目标。

参考文献

- [1] 王燕, 阮绪芝, 孙晓东, 龚坚. 从分子水平理解基因的显性和隐性. *医学理论与实践*, 2011, 24(17): 2059. [DOI]
- [2] Myerowitz R. Tay-Sachs disease-causing mutations and neutral polymorphisms in the Hex A gene. *Human Mutation*, 1997, 9 (3): 195-208. [DOI]
- [3] Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. *Nature*, 1990, 345: 229-233. [DOI]
- [4] Vuorio E, de Crombrughe B. The family of collagen genes. *Annu Rev Biochem*, 1990, 59: 837-872. [DOI]
- [5] Di Lullo GA, Sweeney SM, Korkko J, Ala-Kokko L, San Antonio JD. Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *J Biol Chem*, 2002, 277(6): 4223-4231. [DOI]
- [6] Marini JC, Forlino A, Cabral WA, Barnes AM, San Antonio JD, Milgrom S, Hyland JC, Korkko J, Prockop DJ, De Paepe A, Coucke P, Symoens S, Glorieux FH, Roughley PJ, Lund AM, Kuurila-Svahn K, Hartikka H, Cohn DH, Krakow D, Mottes M, Schwarze U, Chen D, Yang K, Kuslich C, Troendle J, Dagleish R, Byers PH. Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Hum Mutat*, 2007, 28: 209-221. [DOI]
- [7] Sidell BD, O'Brien KM. When bad things happen to good fish: the loss of hemoglobin and myoglobin expression in Antarctic icefishes. *J Exp Biol*, 2006, 209(Pt 10): 1791-1802. [DOI]
- [8] RIGGS A, WELLS M. The oxygen equilibrium of sickle-cell hemoglobin. *Biochim Biophys Acta*, 1961, 50: 243-248. [DOI]
- [9] Harrington DJ, Adachi K, Royer WE Jr. The high resolution crystal structure of deoxyhemoglobin S. *J Mol Biol*, 1997, 272(3): 398-407. [DOI]
- [10] Welles TE, Hayton K, Fairhurst RM. The impact of malaria parasitism: from corpuscles to communities. *J Clin Invest*, 2009, 119 (9): 2496-505. [DOI]
- [11] Makani J, Komba AN, Cox SE, Oruo J, Mwamtemi K, Kitundu J, Magesa P, Rwezaula S, Meda E, Mgaya J, Pallangyo K, Okiro E, Muturi D, Newton CR, Fegan G, Marsh K, Williams TN. Malaria in patients with sickle cell anemia: burden, risk factors, and outcome at the outpatient clinic and during hospitalization. *Blood*, 2010, 115(2): 215-220. [DOI]
- [12] Spyropoulos B. Tay-Sachs carriers and tuberculosis resistance. *Nature*, 1988, 331(6158):666. [DOI]
- [13] Read AP, Strachan T. Human molecular genetics 2. New York: Wiley; 1999. ISBN 0-471-33061-2. Chapter 18: Cancer Genetics. [DOI]
- [14] Halazonetis TD, Kandil AN. Conformational shifts propagate from the oligomerization domain of p53 to its tetrameric DNA binding domain and restore DNA binding to select p53 mutants. *EMBO J*, 1993, 12(13): 5057-5064. [DOI]
- [15] Milner J, Medcalf EA. Cotranslation of activated mutant p53 with wild type drives the wild-type p53 protein into the mutant conformation. *Cell*, 1991, 65(5):765-774. [DOI]
- [16] Zhang AT, Langley AR, Christov CP, Kheir E, Shafee T, Gardiner TJ, Krude T. Dynamic interaction of Y RNAs with chromatin and initiation proteins during human DNA replication. *J Cell Sci*, 2011, 124(Pt 12): 2058-2069. [DOI]
- [17] Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP. The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science*, 2005, 310(5755): 1817-1821. [DOI]
- [18] Meltzer PS. Cancer genomics: small RNAs with big impacts. *Nature*, 2005, 435(7043):745-746. [DOI]
- [19] Finer S, Holland ML, Nanty L, Rakyan VK. The hunt for the epiallele. *Environ Mol Mutagen*, 2011, 52(1): 1-11. [DOI]
- [20] Vrieling H, Duhl DM, Millar SE, Miller KA, Barsh GS. Differences in dorsal and ventral pigmentation result from regional expression of the mouse agouti gene. *Proc Nat Acad Sci USA*, 1994, 91: 5667-5671. [DOI]
- [21] Duhl DM, Stevens ME, Vrieling H, Saxon PJ, Miller MW, Epstein CJ, Barsh GS. Pleiotropic effects of the mouse *le-*

- thal yellow* (A^y) mutation explained by deletion of a maternally expressed gene and the simultaneous production of *agouti* fusion RNAs. *Development*, 1994, 120(6): 1695–1708. [DOI]
- [22] Bultman SJ, Michaud EJ, Woychik RP. Molecular characterization of the mouse *agouti* locus. *Cell*, 1992, 24: 71(7): 1195–204. [DOI]
- [23] Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*, 2003, 72 (1): 291–336. [DOI]
- [24] Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet*, 2008, 40 (12): 1413–1415. [DOI]
- [25] Nembaware V, Lupindo B, Schouest K, Spillane C, Scheffler K, Seoighe C. Genome-wide survey of allele-specific splicing in humans. *BMC Genomics*, 2008, 9: 265. [DOI]
- [26] Zeng L, Fagotto F, Zhang T, Hsu W, Vasicek TJ, Perry WL 3rd, Lee JJ, Tilghman SM, Gumbiner BM, Costantini F. The mouse *Fused* locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell*, 1997, 90(1): 181–192. [DOI]
- [27] Vasicek TJ, Zeng L, Guan XJ, Zhang T, Costantini F, Tilghman SM. Two dominant mutations in the mouse *fused* gene are the result of transposon insertions. *Genetics*, 1997, 147(2):777–786. [DOI]
- [28] Lu Z, Liu W, Huang H, He Y, Han Y, Rui Y, Wang Y, Li Q, Ruan K, Ye Z, Low BC, Meng A, Lin SC. Protein encoded by the Axin(Fu) allele effectively down-regulates Wnt signaling but exerts a dominant negative effect on c-Jun N-terminal kinase signaling. *J Biol Chem*, 2008, 283(19):13132–13139. [DOI]
- [29] Flood WD, Ruvinsky A. Alternative splicing and expressivity of the Axin(Fu) allele in mice. *Heredity* (Edinb). 2001, 87(Pt 2):146–152. [DOI]
- [30] Rakyan VK, Chong S, Champ ME, Cuthbert PC, Morgan HD, Luu KV, Whitelaw E. Transgenerational inheritance of epigenetic states at the murine *Axin^{Fu}* allele occurs after maternal and paternal transmission. *Proc Natl Acad Sci USA*, 2003, 100(5):2538–2543. [DOI]
- [31] Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. doublesex is a mimicry supergene. *Nature*, 2014, 507(7491): 229–32. [DOI]
- [32] Du X, Wang J, Zhu H, Rinaldo L, Lamar KM, Palmenberg AC, Hansel C, Gomez CM. Second cistron in *CACNA1A* gene encodes a transcription factor mediating cerebellar development and SCA6. *Cell*, 2013, 154:118–133. [DOI]
- [33] Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. *Trends Genet*, 2002, 18: 348–351. [DOI]
- [34] Rakyan VK, Preis J, Morgan HD, Whitelaw E. The marks, mechanisms and memory of epigenetic states in mammals. *Biochem J*, 2001, 356: 1–10. [DOI]
- [35] Dolinoy DC, Weinhouse C, Jones TR, Rozek LS, Jirtle RL. Variable histone modifications at the *Avy* metastable epiallele. *Epigenetics*, 2010, 5(7): 637–644. [DOI]
- [36] Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol*, 2003, 23(15): 5293–5300. [DOI]
- [37] Cockett NE, Jackson SP, Shay TL, Nielsen D, Moore SS, Steele MR, Barendse W, Green RD, Georges M. Chromosomal localisation of the callipyge gene in sheep (*Ovis aries*) using bovine DNA markers. *Proc Natl Acad Sci USA*, 1994, 91(8): 3019–3023. [DOI]
- [38] Cockett NE, Jackson SP, Shay TL, Farnir F, Berghmans S, Snowden GD, Nielsen DM, Georges M. Polar overdominance at the ovine callipyge locus. *Science*, 1996, 273(5272): 236–238. [DOI]
- [39] Freking BA, Murphy SK, Wylie AA, Rhodes SJ, Keele JW, Leymaster KA, Jirtle RL, Smith TP. Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. *Genome Res*, 2002, 12(10): 1496–1506. [DOI]
- [40] Smit M, Segers K, Carrascosa LG, Shay T, Baraldi F, Gyapay G, Snowden G, Georges M, Cockett N, Charlier C. Mosaicism of Solid Gold supports the causality of a non-coding A to G transition in the determinism of the callipyge phenotype. *Genetics*, 2003, 163(1): 453–456. [DOI]
- [41] Davis E, Jensen CH, Schroder HD, Farnir F, Shay-Hadfield T, Kliem A, Cockett N, Georges M, Charlier C. Ectopic expression of DLK1 protein in skeletal muscle of padumal heterozygotes causes the callipyge phenotype. *Curr Biol*, 2004, 14(20): 1858–1862. [DOI]
- [42] Georges M, Charlier C, Smit M, Davis E, Shay T, Tordoir X, Takeda H, Caiment F, Cockett N. Toward molecular understanding of polar overdominance at the ovine callipyge locus. *Cold Spring Harb Symp Quant Biol*, 2004, 69: 477–483. [DOI]
- [43] Murphy SK, Nolan CM, Huang Z, Kucera KS, Freking BA, Smith TP, Leymaster KA, Weidman JR, Jirtle RL. Callipyge mutation affects gene expression in cis: a potential role for chromatin structure. *Genome Res*, 2006, 16(3): 340–346. [DOI]
- [44] Takeda H, Caiment F, Smit M, Hiard S, Tordoir X, Cockett N, Georges M, Charlier C. The callipyge mutation enhances bidirectional long-range DLK1-GTL2 intergenic transcription in cis. *Proc Natl Acad Sci USA*, 2006, 103(21): 8119–8124. [DOI]
- [45] De Gregorio L, Jinnah HA, Harris JC, Nyhan WL, Schretlen DJ, Trombley LM, O'Neill JP. Lesch-Nyhan disease in a female with a clinically normal monozygotic twin. *Mol Genet Metab*, 2005, 85: 70–77. [DOI]