7:38 🙎 夸 📶 💶 56% \leftarrow 137 | 问答: 人和黑猩猩 的基因只有1%差异? 卓克 昨天 137 | 问答: 人和黑猩猩的... 08:23 3.92MB |卓克亲述| 这周我留了一个稍微复杂一些的问题,都 说猩猩跟人在基因上高度类似,有99%都 类似,这个真假对错很多人是拿不准的。 但是如果你细心的话就会发现,往往还有 这样的后续内容,就是说猪跟人基因的相 似度是90%,实验中小老鼠的基因跟人类 相似度都有80%。 好,确实,猪比小老鼠更像人一些,所以 相似度一个是90%,一个是80%。我们也 默认这个是对的,矛盾并没有出现。 但是我们还能见到有一些介绍遗传病的科 普里是这么说的: 异卵双胞胎拥有50%相同的基因,因为 兄弟间每个人都各自从父母那儿取得了 一半的染色体,相当于两个人各做了46 次二选一, 所以最终总体上看来, 大约 有50%是相同的。 这下矛盾来了,难道我们跟小老鼠的相似 度要远高于我跟我哥哥吗?这个问题出在 哪呢? 就在于定义不清晰的词语和不加质疑地相 信传言。 人跟黑猩猩有99%的相似度,这种说法确 实不是瞎编的,是来自于1975年《科学》 杂志上的一篇对比人跟黑猩猩体内大约43 种蛋白质的相似度的这么一篇论文。 SCIENCE 11 April 1975, Volume 188, Number 4184 evidence concerning the molecular basis of evolution at the organismal level. We suggest that evolutionary changes in anatomy and way of life are more often based on changes in the mecha-**Evolution at Two Levels in** nisms controlling the expression of genes than on sequence changes in pro-**Humans and Chimpanzees** teins. We therefore propose that regulatory mutations account for the major biological differences between humans Their macromolecules are so alike that regulatory and chimpanzees. mutations may account for their biological differences. Similarity of Human and Mary-Claire King and A. C. Wilson Chimpanzee Genes To compare human and chimpanzee genes, one compares either homologous proteins or nucleic acids. At the protein (Pan troglodytes) and humans (Homo Soon after the expansion of molecular level, one way of measuring the degree biology in the 1950's, it became evident sapiens). This pair of species is also of genetic similarity of two taxa is to because of the thoroughness by comparing the proteins and determine the average number of amino nucleic acids of one species with those of another, one could hope to obtain with which they have been compared acid differences between homologous at the organismal level-that is, at the polypeptides from each population. The a quantitative and objective estimate level of anatomy, physiology, behavior, most direct method for determining this 论文截图 这篇论文最终的结论就是: 人跟黑猩猩体内多肽的检测发现,平均 来说,有99%以上都是相同的。 多肽是什么呢? 其实就是蛋白质的雏形,多肽连在一起折 叠好,就是蛋白质了。 但是我们完全不能把刚刚那个结论理解成 人 跟 黑 猩 猩 体 内 的 蛋 白 质 的 相 似 度 是 因为在这篇论文中,只比较了40多 个多肽。 可人体究竟有多少种蛋白质,多少种多肽

呢?

有多高呢?

都是2005年。

那怎么比呢?

就发现:

体的内容。

子"来作对比的。

什么是转座子呢?

变的。

件。

现在还有争论,但起码种类不会低于10万

种。我们对于人的了解都是如此,我们就

更不知道黑猩猩体内到底有多少种蛋白质

了。所以这项研究,当时只是一个小小的

尝试,只是充分利用了那个年代刚刚诞生

从科学技术进展的角度来看,人类最早在

什么时候才能说人跟黑猩猩的基因相似度

至少也要等到人类基因组跟黑猩猩基因组

全部测序完才行,而这两个时间点差不多

比如说先从承载基因的染色体入手,结果

人有23对染色体,黑猩猩有24对,这个相

其实就有点像两本书的相似度的比较,比

如说对比1990年高二的数学课本跟2010

年高二的数学课本,你就发现90年版的一

共23章,10年版的一共24章,这就不一

样。那你说我不管多少章,也许只是因为

重新打散了、重新排的原因,所以章数不

一样,它们的内容也许还是高度一致的。

好,那咱们就不管多少章,咱们只对比具

在基因相似度的对比中,就有选"转座

通俗来说就是一段基因,它可以到处地挪

动位置,但是这段基因的序列内部是不会

我们还用教科书来举例子,转座子就有点

类似于像概率论这样的知识,比如说从5

个小球里选出2个,有几种可能的方案

呢?这部分知识你放在高三学也可以,你

放在初二学也行,甚至小学三年级也可

以,要学懂这块知识并不需要什么前置条

那么我们就统计,在两个版本的教科书

里,这种类型的整块的知识分别有多少

比如像 Alu(基因中特征明显的片段) 这

个知识点,黑猩猩有2340个,人有7000

多个,在这个知识点的数据量,黑猩猩是

0.7兆,人的数据量是2.1兆,你看这就有

很大区别。还有像小卫星的片段(基因中

特征明显的片段),黑猩猩有7000多个,

人类有11000多个,这样不同的转座子非

那么在黑猩猩所有的序列中,凡是可以和

人匹配的都给它算上,这些数据量占比人

所以你看,在这种对比方法下,所谓的

DNA 相似性是有它自己的定义的,而且还

是比较明确的。就是先在黑猩猩的序列

中,把所有人跟黑猩猩都有的序列挑出

来,然后再除以人类的数据总量,就是相

其实也是有很多弹性空间的。比如说有这

么一段基因,它仅仅由3个碱基组成,当

然实际情况中不可能这么少,可能至少都

得有3000,30000个,那咱们就用3个来

举例:人是 CCC,黑猩猩是 CCA,狗是

那可能在某一种判断方法里,会简单粗暴

地认为,人、黑猩猩、狗都拥有这个基

因,只不过是基因的多肽性造成的差异,

这不能算作不匹配,所以,人、黑猩猩、

狗在这个基因的相似度上是100%,而猪

因为没有这段基因, 所以猪跟人在这段基

因上的相似度是0。而在另外的评价标准

中,黑猩猩跟人的这段序列相差了1/3,

所以即便是细致地比较基因序列,相似度

的弹性空间也是比较大的。现在的研究结

果,人跟黑猩猩在基因序列上相似度是

72%到89%之间,不论它是大是小,都肯

知道这个道理,我们就可以回想,有些人

介绍人类祖先的内容的时候就会说到,人

类到现在还有4%左右的 DNA 源自于尼安

就是因为当年在测试尼安德特人的 DNA

的时候并不轻松,都是从一些化石上,比

如化石的小指骨上取一些材料,来测定这

份已经碳化的骨骼里的 DNA 序列,这里

当然可能有尼安德特人的 DNA, 那还有可

能有其他动植物,或者是微生物的 DNA。

人类学家在研究的时候一定会尽量地排除

其他干扰因素,但最终确定是尼安德特人

的 DNA,这个序列也不可能是全基因组的

序列,也许只是一个尼安德特人全序列的

5%, 10%啊, 那么把这个已经测定好的序

列作为模板,拿到人类的 DNA 序列中去

碰,看看和这一小部分相同的序列能够占

到人类 DNA 序列总量的多少呢?这样就

能得出一个数据。比如说,人类至少还包

其实,我们从道理上也能明白,我们在遗

传距离上一定是离尼安德特人更近,离黑

猩猩更远,数据完备的情况下,黑猩猩和

我们至少都有72%是相同的,那我们跟尼

所以你看,一个问题要有像样的答案,真

要仔细琢磨问题中出现的每一个词语的含

义,这个思考习惯咱们专栏的用户应该已

经慢慢地有意识地建立起来了。在跟朋友

聊天谈话的时候,虽然没必要较真,但真

到了一定要说出个123的时候,还是免不

跟黑猩猩有99%的相似度,说的

添加到知识账本

请朋友读

了要弄清所有的概念。

■ 划重点

Aа

安德特人的相似度至少会高于这个数字。

含有4%的尼安德特人的 DNA。

定不会有98%或99%这么高。

那这句话是怎么来的呢?

德特人。

所以相似性是66%,狗是33%,猪是0。

CAA,猪没有这个基因。

什么能算作匹配,什么不能算匹配

类基因组数据的总量大约是多少呢?

最终的结果是75%左右。

似性了。

呢?

常非常多,它们都属于不匹配的。

个,每个知识点的字数是多少。

似度的比较有点像什么呢?

的,确定氨基酸序列方法的一种小尝试。