[](https://www.frontiersin.org/journals/plant-science)[类型原创研究2022年9月21日公布](https://www.frontiersin.org/journals/plant-science#editorial-board)[DOI 10.3389/fpls . 2022.96976565666](https://doi.org/10.3389/fpls.2022.969765)



[](http://crossmark.crossref.org/dialog/?doi=10.3389/fpls.2022.969765&domain=pdf&date_stamp=2022-09-21)

开架借阅

编辑人

汤春正，

中国北京林业大学

由...审查

日本岛根大学小林信夫，

中国杨勇贵州大学，

中国浙江农业科学院

\*通信

金松亨[shjin@zafu.edu.cn](mailto:shjin@zafu.edu.cn)

专业部分

本文提交给植物生物信息学，

植物科学前沿杂志的一部分

2022年6月15日收到

2022年8月30日接受

2022年9月21日公布

引用

沈杰，李x，李m，程h，黄x，金S (2022)杜鹃花细胞器基因组的特征、比较系统发育和基因转移分析。

×

*前面。植物科学。13:969765.doi:*[10.3389/fpls.2022.969765](https://doi.org/10.3389/fpls.2022.969765)

版权

沈，李，李，程，黄，金。这是一篇根据[Creative Commons 属性License(CC BY)](http://creativecommons.org/licenses/by/4.0/)。根据公认的学术惯例，允许在其他论坛上使用、分发或复制，但前提是原创作者和版权所有者获得认可，并且本期刊的原创出版物被引用。不允许使用、分发或复制不符合这些条款的内容。

[杜鹃花细胞器基因组的特征、比较系统发育和基因转移分析](https://www.frontiersin.org/articles/10.3389/fpls.2022.969765/full)**×**[*pulchrum*](https://www.frontiersin.org/articles/10.3389/fpls.2022.969765/full)

建爽身1、雪芹历1、2、智明李、鹤峰成1、小令黄1、宋衡金1、2\*

1浙江A&F大学济阳学院，浙江诸暨，2生命科学与健康部，湖州学院，浙江湖州，3生物与数据生物技术有限公司，广州

杜鹃花是一种重要的园艺植物，广泛分布于欧洲、亚洲和北美。为分析锦绣杜鹃及其近缘种的系统发育和细胞器基因组信息，对锦绣杜鹃的细胞器基因组进行了测序和组装。线粒体全基因组为谱系DNA分子，长816，410 bp，含64个基因，即24个转移RNA (tRNA)基因、3个核糖体RNA (rRNA)基因和37个蛋白编码基因。重组了R. ×普勒姆叶绿体基因组并重新注释；结果与以往的研究不同。从菹草线粒体和叶绿体基因组中分别鉴定出42个和46个简单序列重复序列(SSR)。五个基因(nad1、nad2、nad4、nad7和rps3)是潜在的有用分子标记。杜鹃花科5种植物的R. ×普勒姆线粒体基因组共线排列显示，这些相关物种的线粒体基因组在该基因区域与R. ×普勒姆具有高度同源性，最保守的基因为trnC-GCA、trnD- GUC、trnM-CAU、trnN-GUU、trnY-GUA、atp4、nad4、nad2、nad5、ccmC和rrn26。线粒体基因组进化树结果表明，西门塔尔牛是团头鲂的姊妹种。结果证实了普勒姆和西门塔尔牛线粒体基因组之间存在基因重排。10个杜鹃花科线粒体基因和

7个杜鹃花叶绿体基因受到突变的影响，而其他基因的密码子使用都经历了选择。本研究共鉴定出13个同源片段，其中包含R. ×普勒姆叶绿体和线粒体基因组间的基因序列。总的来说，我们的结果表明

细胞器基因组信息可以解释植物的系统发育

并且可以用于开发分子标记和遗传进化。我们的研究将有助于杜鹃花和杜鹃花科其他属的种群遗传学和进化研究。

关键词

*杜鹃花，线粒体基因组，叶绿体基因组，基因转移，系统发育*

# 介绍

植物含有两种细胞器:质体和线粒体。这些细胞器保留了它们自己的基因组，这些基因组独立于核基因组([Gray](#_bookmark28),[2012](#_bookmark28)).高等植物的质体基因组在大小、基因含量、基因结构和基因顺序上是保守的([Sugiura](#_bookmark65),[1995](#_bookmark65)；[Wicke et al.](#_bookmark74),[2011](#_bookmark74)).大多数质体是由双链DNA组成的环状分子，含有4个典型区域，大小从72到217 kb不等，并含有130个基因([Jansen et al.](#_bookmark34),[2011](#_bookmark34)).具有保守结构和高取代率的质体基因组用于研究植物的系统发育、生物学和光合基因降解([Wu and Ge](#_bookmark76),[2012](#_bookmark76)；[Thomas et al.](#_bookmark68),[2017](#_bookmark68)；[Shen et al.](#_bookmark60),[2020](#_bookmark60)).与质体基因组相比，植物线粒体基因组具有分布广泛、大小从66到11，000 kb、多部分基因组变异、种间排列、基因序列转移或缺失等独特的特征([Knoop](#_bookmark37),[2004](#_bookmark37)；[Parkinson et al.](#_bookmark49),[2005](#_bookmark49)；[Sloanet al.](#_bookmark62),[2012](#_bookmark62)；[Wu et al.](#_bookmark77),[2022](#_bookmark77)).已在植物中探索了线粒体基因组大小、染色体数目和拷贝数变异([Alverson et al.](#_bookmark15),[2011](#_bookmark15)).例如，拟南芥线粒体基因组具有典型的圆形结构([Sloan et al.](#_bookmark63),[2018](#_bookmark63))，而在Silene圆锥形和藜形中，线粒体基因组具有线性或多染色体结构([Backert and Borner](#_bookmark17),[2000](#_bookmark17)；[Sloan et al.](#_bookmark62),[2012](#_bookmark62)；[Sanchez-Puertaet al.](#_bookmark53),[2017](#_bookmark53)).线粒体基因组可用于开发分子标记和分析线粒体基因组扩增机制([Zubaer et al.](#_bookmark81),[2018](#_bookmark81)).将线粒体基因组与相关物种进行比较可以为解释进化机制和线粒体基因组重排以及鉴定物种分类群提供新的途径([Li et al.](#_bookmark43),

[2018](#_bookmark43)；[Zubaer et al.](#_bookmark81),[2018](#_bookmark81)).

在Illumina Hiseq平台上对植物的整个细胞器基因组进行了测序([Shen et al.](#_bookmark60),[2020](#_bookmark60)；[Xu et al.](#_bookmark79),[2021](#_bookmark79)).然而，Illumina读取长度通常不会跨越更长的重复；这些区域不完全装配，因此影响基因组长度和内容的准确性。随着这一技术的发展，第三代测序(TGS)方法，如牛津纳米孔和长阅读长度的PacBio测序，可以提高以前未装配的基因组区域的覆盖率和装配精度，是理解植物细胞器基因组信息的有用工具([Shearman et al.](#_bookmark58),[2016](#_bookmark58)).目前有数以千计的植物质体基因组和数百个完整的陆地植物线粒体基因组，其中大部分来自作物物种([Wu et al.](#_bookmark77),[2022](#_bookmark77)).杜鹃花目（Ericales）由大约25个科组成([Anderberg et al.](#_bookmark16),[2002](#_bookmark16))，只有5科10种植物的线粒体基因组数据发表在NCBI GenBank数据库中。杜鹃属的线粒体基因组很少发表，仅报道简单的基因注释信息

（[Xu et al.](#_bookmark79),[2021](#_bookmark79)).线粒体基因组中的基因可以为杜鹃花属物种间的系统发育关系提供新的解释。需要使用从多个基因中积累的分子数据进一步研究杜鹃花科的系统发育相互关系([Anderberg et al.](#_bookmark16),[2002](#_bookmark16)).目前，NCBI基因库数据库中有杜鹃花科（Ericaceae）杜鹃花属11种植物的完整质体序列；这些包括五个物种(德拉瓦伊杜鹃变种。Delavayi，R. griersonianum，R. henanense亚种。lingbaoense，R. platypodum和R. delavayi。)属于亚门。处女膜动物，一种(R. molle)属于亚属。五花草，以及属于亚属的五种(R. kawakamii，R. datiandingense，R. micranthum，R. concinnum和R. simsii)。Tsutsusi。这些杜鹃花叶绿体基因组大部分是通过第二代测序组装的，分析了基因组结构和系统发育([Thomas et al.](#_bookmark68),[2017](#_bookmark68)；[Li et al.](#_bookmark43),[2018](#_bookmark43)；[Zubaer et al.](#_bookmark81),[2018](#_bookmark81)).这些研究的结果大不相同。如是否退出IRs区域，结果令人惊讶的是不同Rhodordendron物种之间的差异。与质体序列相比，依赖于有限的基因组数据，人们对Rhodordendron物种线粒体基因组相关的基因组变异、排列、叶绿体到线粒体的基因转移研究了解甚少。随着新技术的发展，越来越多的细胞器基因组数据被发表和分析，叶绿体到线粒体的基因转移被认为是长期进化的特征([Gui et al.](#_bookmark30),[2016](#_bookmark30)；[Nguyen et al.](#_bookmark47),[2020](#_bookmark47)).以往的研究主要集中在被子植物细胞器中细胞核DNA的基因转移([Smith](#_bookmark64),[2011](#_bookmark64)；[Park et al.](#_bookmark48),[2014](#_bookmark48)).因此，对红顶蕨属物种线粒体和质体基因组信息的研究还有待进一步深入。

*锦绣杜鹃是一种重要的园艺植物，用途广泛*分布在欧洲、亚洲和北美的温带地区([Galle](#_bookmark26),[1985](#_bookmark26)).r×pul chrum被认为是“Omurasaki”和平度杜鹃花品种的园艺品种。最近的研究揭示了*R.××普勒姆及其具有推定祖先种的相关栽培品种(包括毛萼木、巨萼木、*

*R.使用F3"5'H基因序列和AFLP技术的糙皮病菌、R. ×普勒姆"村崎"和R. ×短翅病菌" Shiro-rykye "(*[Scariot et al.](#_bookmark55),[2007](#_bookmark55)；[Meanchaipiboon et al.](#_bookmark46),[2021](#_bookmark46)).此外，在中还报告了“Omurasaki”和平度杜鹃花品种群的叶绿体(cpDNA)起源[Kobayashi](#_bookmark38)

[et al.](#_bookmark38)（[2021](#_bookmark38)).“Omurasaki”和大部分平度杜鹃花栽培品种拥有日本野生杜鹃花的cpDNA([Shirasawa et al.](#_bookmark61),[2021](#_bookmark61)).到目前为止，还没有关于锦绣杜鹃线粒体基因组的报道。需要进一步研究杜鹃属物种的线粒体和质体基因组信息，这将有助于杜鹃种群遗传学和进化的研究

有助于了解杜鹃花科植物的进化机制和物种分类群的鉴定。使用Illumina Hiseq平台对之前发表的R. × pulchrum质体基因组进行测序([Shen et al.](#_bookmark59),[2019](#_bookmark59),[2020](#_bookmark60)).本研究采用TGS方法结合第二代测序技术，检测了菹草的完整线粒体和质体基因组。我们组装并注释了的完整线粒体和质体基因组

*R.××白粉菌并分析了其基因组的含量、组织、*

和系统发育分析。我们对杜鹃属物种进行了比较有丝分裂基因组分析，以确定基因组中的变异、保守和重排区域。我们还分析了r××pul chrum线粒体基因组和质体基因组之间的基因转移。线粒体和质体基因组信息可以解释植物的系统发育和进化关系，可用于开发分子标记和基因工程。

# 材料和方法

## DNA提取和测序

紫色大花型杜鹃花×白花杜鹃品种扦插无性系的嫩绿叶([**Supplementary Figure 1**](#_bookmark14))采自浙江A&F大学苗圃(保藏于中国科学院植物研究所Mem，标本保藏号为PE00820836)并立即在-80℃保藏。使用改良的十六烷基三甲基溴化铵(CTAB)法从幼叶中提取总基因组DNA([Doyle](#_bookmark24),[1987](#_bookmark24)).高质量的DNA被用于随后的文库制备，并使用promex和BGISEQ-500平台(Bio & Data Biotechnologies Co .，广州，中国)进行测序。为了获得长的非片段序列reads，剪切了15 g的基因组DNA，并用BluePippin (Sage Science，Beverly，MA，USA)进行大小选择(30–80kb)。按照制造商的说明，使用连接测序1D试剂盒(牛津纳米孔，牛津，英国)处理选定的片段，并使用PromethION DNA测序仪(牛津纳米孔，牛津，英国)测序48–72h。

DNA提取后，我们将1 g纯化的DNA片段化，并使用它建立了300 bp的短插入文库。根据制造商的说明，在bgi eq-500测序仪上用PE150 bp对这些合格的文库进行了测序。测序采用黑桃v-3.13.0软件([Bankevich et al.](#_bookmark18),[2012](#_bookmark18))，使用Canu v-1.5软件对TGS数据进行单独汇编([Koren et al.](#_bookmark39),[2017](#_bookmark39)).

## 顺序读取的预处理

对于长reads，使用Porechop v0.2.4进行适配器微调[1](#_bookmark0)并且使用Guppy进行质量得分< 7的删除读取。对于简短读取，原始读取由Fastp v.0.20.1用默认参数([Chenet al.](#_bookmark21),[2018](#_bookmark21))以修整适配器并去除低质量reads(Phred质量分数< 20)。

## 线粒体基因组组装

使用SPAdes v-3.13.0软件(参数−k 21,33,55,77,89 –careful,其他默认)进行第三代和第二代测序的混合从头组装([Bankevich et al.](#_bookmark18),[2012](#_bookmark18))，使用Canu v-1.5软件单独组装第三代(长)测序数据，参数设置为(1)基因组大小0.8 Mb和(2)校正错误率= 0.03([Koren et al.](#_bookmark39),[2017](#_bookmark39)).将组装的contigs与NCBI杜鹃花目物种的所有线粒体序列进行比对，并提取候选线粒体contigs。然后使用Pilon软件抛光候选线粒体contigs([Walker et al.](#_bookmark71),[2014](#_bookmark71))，使用Geneious prime软件对contigs进行第二代(短)测序reads延伸([Kearse et al.](#_bookmark36),[2012](#_bookmark36)).使用Geneious Prime软件鉴定所选线粒体contigs末端的重复序列。基于末端重复序列连接contigs，对组装的长contigs进行短reads比较和末端延伸，直到没有reads可以用来进一步延伸([Kearse et al.](#_bookmark36),[2012](#_bookmark36)).通过比较4种杜鹃花目植物(即西门塔尔牛(NC053763)、大果越橘(NC023338)、低三苯基单孢菌(MK990822)和蓝萼龙舌兰(MZ779111))的线粒体基因组作为参考来确认组装。

## 叶绿体基因组的组装

使用Minimap2软件将TGSreads与来自NCBI的杜鹃花属物种的所有叶绿体基因组数据进行比对([Li](#_bookmark41),[2018](#_bookmark41))，并提取比对长度大于5，000 bp的reads用于随后的组装。从Genbank下载第二代测序reads，并使用Bowtie2软件([Langmead and Salzberg](#_bookmark40),[2012](#_bookmark40))，对齐的reads用于随后的组装。叶绿体候选物第三代和第二代读取

组装的contigs与所有制造商的contigs对齐

说明书详见文献([Huang et al.](#_bookmark33),[2017](#_bookmark33)).

一<https://github.com/rrwick/Porechop>

使用Unicycler版本:v0.4.8软件(带有默认参数([Wick et al.](#_bookmark73),[2017](#_bookmark73)).本研究重组并重新注释了R. ×普勒姆叶绿体基因组；所有新读取的内容都保存在NCBI序列读取档案(SRA)中，注册号为MN182619.2。

已使用LASTZ软件(版本1.02.00)比较了杜鹃×黄花杜鹃的MN182619.1和MN182619.2之间比较的完整叶绿体基因组共线比对([Harris](#_bookmark31),[2007](#_bookmark31)；[Kearse et al.](#_bookmark36),[2012](#_bookmark36)).随机选取断裂和倒位位点设计聚合酶链式反应(PCR)引物，验证叶绿体基因组的准确性。

## 线粒体和叶绿体基因组注释

线粒体和叶绿体基因组注释采用GeSeq软件([Tillich et al.](#_bookmark69),[2017](#_bookmark69))，并使用Geneious prime手动更正注释结果([Kearse et al.](#_bookmark36),[2012](#_bookmark36)).使用细胞器基因组绘图工具(OGDRAW) v.1.3.1绘制了基因组图，以进一步比较基因顺序和含量([Greiner et al.](#_bookmark29),[2019](#_bookmark29)).相对同义密码子使用量(RSCU)计算如下[Sharp and Li](#_bookmark57) （[1986](#_bookmark57)).

## 重复序列的识别

使用MISA(微卫星鉴定工具)软件进行SSR分析的简单序列重复(SSR)([Thiel et al.](#_bookmark67),[2003](#_bookmark67))，参数设置如下(unit\_size，min \_ repeats):1–10 2–6 3–5 4–5 5 6–5，中断(max \_ difference \_ by \_ 2 \_ SSRs):100。使用串联重复检测器分析串联重复序列，参数设置为2、7、7、80、10、50、500、f、d和m([Benson](#_bookmark19),[1999](#_bookmark19)).

## 系统发育分析

系统发育采用最大似然(ML)法，使用Fasttree软件([Price et al.](#_bookmark51),[2010](#_bookmark51)).最初使用MAFFTv7.313对序列进行比对([Katoh and Standley](#_bookmark35),[2013](#_bookmark35))，并使用Fasttree 2软件在GTR +伽玛模型下构建ML树([Price et al.](#_bookmark51),[2010](#_bookmark51)).

线粒体系统发育树选择了13个相关物种的线粒体基因组共有的8个编码基因片段(atp1、atp4、atp9、ccmC、matR、nad3、nad6和rps12):

*R.ápul chrum(om 283814)、R. simsii (NC053763)、痘苗*

*山茱萸(NC023338)、绿盲蝽(MZ779111)、单绿盲蝽(MK990822)、桐花树(NC056358)、棘胸蛙(MZ151883)、茶树*

var。Assamica (MK574877)，茶树(NC043914)，

*中华猕猴桃(MZ959063)、中华猕猴桃(MZ959061)和*

*A.argusta(MH 559343)。使用葡萄(NC012119)作为外组样本进行序列比对。本研究构建了杜鹃花属线粒体进化树。13个相关物种中8个基因的同源部分的长度见*[**Supplementary Table 1**](#_bookmark14)。

叶绿体进化树选择了57个编码基因片段(atpA、atpB、atpE、atpH、atpI、cemA、matK、ndhA、ndhB、ndhC、ndhD、ndhE、ndhF、ndhH、ndhI、ndhJ、petA、petB、petD、petG、petL、petN、psaA、psaB、psaC、psaI、psbA、psbB、psbC、psbD、psbE、psbF、psbJ、psbK、psbL、psbM、delavayi (NC047438)、r . griersonium(NC 050162)、R. henanense亚种。lingboaense(mt 239363)、r . plate podum(mt 985162)、R. delavayi (MN711645)、R. molle (MZ073672)、r . kawakami(NC 058233)，

*R.datiandingenase(NC 057644)，R. micranthum (MT239365)，r . concent um(m t 239366)，R. ripense (DRR298903-DRR298907)，*

*R.卵形体(SRR12917131-SRR12917132)、r . latouchae(SRR 13425299)、r . x pul chrum(Mn 182619.2)和R. simsii (MT239364)。美味猕猴桃(NC026691)和中国猕猴桃(NC026690)被用作组外样本进行序列比对。杜鹃花属具有广泛的叶绿体数据；因此，本研究仅构建了杜鹃花叶绿体进化树。15个相关物种的57个基因的同源部分的长度已在*[**Supplementary Table 2**](#_bookmark14)。

## 密码子使用偏向模式分析

密码子使用偏性参数RSCU和有效密码子数(ENC)采用CondonW 1.4.4软件计算([Peden and John](#_bookmark50),[2000](#_bookmark50)).ENC用于分析基因碱基组成对密码子使用偏性的影响([Wright](#_bookmark75),[1990](#_bookmark75)).这两个参数被用来描述密码子使用偏好的模式。从杜鹃花科植物中筛选出5个线粒体基因组，用于密码子使用偏倚参数的计算，包括单孢原毛霉(MK990822)、蓝萼龙葵(MZ779111)、大果越橘(NC023338)、西门塔尔牛(NC053763)和r×普勒姆(OM283814)。选择15个杜鹃花叶绿体基因组计算密码子使用偏倚参数，包括德氏杜鹃花。delavayi (NC047438)、r . griersonium(NC 050162)、R. henanense亚种。lingboaense(mt 239363)、r . plate podum(mt 985162)、R. delavayi (MN711645)、R. molle (MZ073672)、r . kawakami(NC 058233)，

*R.datiandingenase(NC 057644)，R. micranthum (MT239365)，r . concent um(m t 239366)，R. ripense (DRR298903-DRR298907)，*

*R.卵形体(SRR12917131-SRR12917132)、拉图科r(SRR 13425299)和西门塔尔牛(MT239364)以及r×普勒姆(MN182619.2)。*

## 近缘种的基因组比较及叶绿体和线粒体基因组间的水平基因转移

以r×pul chrum线粒体基因组为对照，与西门塔尔牛(NC\_053763)比较，测定线粒体重排。采用Mauve软件分析线粒体重排，绘制结构变异图([Darling et al.](#_bookmark22),[2004](#_bookmark22)).

中华绒螯蟹线粒体基因组(OM283814)

和R. simsii (NC\_053763)进行共线性分析([Rice et al.](#_bookmark52),[2000](#_bookmark52)).LASTZ版本

1.02.00用于进行全基因共线比对以检测共线片段([Harris](#_bookmark31),[2007](#_bookmark31)；[Kearse et al.](#_bookmark36),[2012](#_bookmark36)).参数设置如下:step = 20和

种子模式= 12/19。使用显示结果

以中华绒螯蟹线粒体基因组为对照

与属于杜鹃花科的其他四个基因组(西门塔尔牛、山茱萸、绿盲蝽和发育不全分枝杆菌)相比。叶绿体和线粒体基因组之间的水平基因转移(HGT)是使用LASTZ版本确定的

1.02.00 ([Harris](#_bookmark31),[2007](#_bookmark31)).

# 结果

## 线粒体基因组组织和特征

在本研究中，线粒体基因组组装的平均阅读覆盖率高于325×104。R. ×普勒姆线粒体基因组的基因库登录号为OM283814。全线粒体基因组(有丝分裂基因组)测序从第二代测序平台产生了36，533，204个干净reads(SRA注册号SRR17758703)和5，364，975个通过reads(SRA

总共注释了9个含内含子的基因(rps14、Cox2、ccmFC、rps3、nad1、nad 2、nad4、nad5和nad7)([**Table 1**](#_bookmark1))，其中nad4有2个内含子的蛋白编码基因，nad1、nad 2、nad5、nad7蛋白编码基因有4个内含子，其余有1个内含子。

然后，我们根据蛋白编码基因和tRNA基因估计密码子使用频率。如所示[**SupplementaryTable 5**](#_bookmark14) 和[**Supplementary Figure 2**](#_bookmark14)、UUC、AUU和UUU是最常见的密码子，导致R. × pulchrum线粒体基因组的AU含量高(54.3%)。线粒体基因组由11，196个密码子(65种不同类型)组成，编码20个氨基酸，其中亮氨酸(Leu)是最常用的氨基酸(11.59%，1，298)，半胱氨酸(Cys)最少(1.46%，163)()[**Supplementary Table 5**](#_bookmark14)).结果表明

*R.××蒲黄线粒体基因组偏好同义*

以A结尾的密码子或以RSCU > 1结尾的密码子，但以g结尾的UUG密码子除外。

从杜鹃花中鉴定出42个SSR标记

线粒体基因组，36个单核苷酸(63.35%)，

4个二核苷酸(2.71%)，2个三核苷酸(26.24%)([**Supplementary Table 6**](#_bookmark14)).两个SSR含有鸟嘌呤(G)或胞嘧啶(C)，而其余34个SSR含有聚腺嘌呤(poly A)或聚乙胺(poly T)。6个SSR标记位于包括AG在内的基因编码区

(6)和CT (6)位于nad4基因，两个A (10)位于nad1和nad2基因，两个A (12)位于nad7和rps3基因。在r×/pul chrum线粒体基因组中共鉴定出28个串联重复序列(TR)

表1白花杜鹃注释基因

**×**

线粒体基因组。

### 基因组基因的名称

核糖体RNA rrn5，rrn18，rrn26

转移RNA trnC-GCA、trnD-GUC、trnE-UUC2、trnF-GAA、trnG-GCC、trnH-GUG、trnI-CAU、trnK-UUU、trnM-CAU3、trnN-GUU、trnP-UGG、

*trnQ-UUG，trnS-UGA，trnS-GCU1，*

*trnW-CCA，trnY-GUA，*

核糖体小亚基rps1、rps144、rps4、rpl5、rpl10、rps12、rps144，ATP合酶的亚基atp1、atp4、atp6、atp8、atp91

注册号SRR17758704)，来自TGS平台。中汇总了排序和组装统计信息[**Supplementary Table 3**](#_bookmark14)。

线粒体基因组由谱系DNA分子组成，在816，410 bp处组装，具有GC含量

NADH脱氢酶(复合物I)

细胞色素c氧化酶(复合物IV)

育碧醇细胞色素c还原酶(复合物ⅲ)

*nad16，nad26，nad3，nad41，5，nad4L，nad56，*

*nad6，nad76，nad9 cox11，cox24，cox3，*

*雄天鹅*

锦绣杜鹃为45.7%([**Figure 1**](#_bookmark2)和[**Supplementary**](#_bookmark14)

Maturases

[**Table 4**](#_bookmark14)).锦绣杜鹃线粒体全基因组含有64个注释基因，即24个转移RNA (tRNA)

琥珀酸脱氢酶(复合物ⅱ)

*sdh3、sdh4*

基因、3个核糖体RNA (rRNA)基因和37个蛋白编码基因(mRNA)([**Table 1**](#_bookmark1)).这些基因的位置和排列如所示[**Figure 1**](#_bookmark2)。

其他基因ccmB、ccmC、ccmFC4、ccmFN、mttB、ycf68

1两个基因拷贝；2 .三个基因拷贝；3 . 6个基因拷贝；4个含有单个内含子的基因；

5个含有两个内含子的基因；6含有两个内含子的基因。

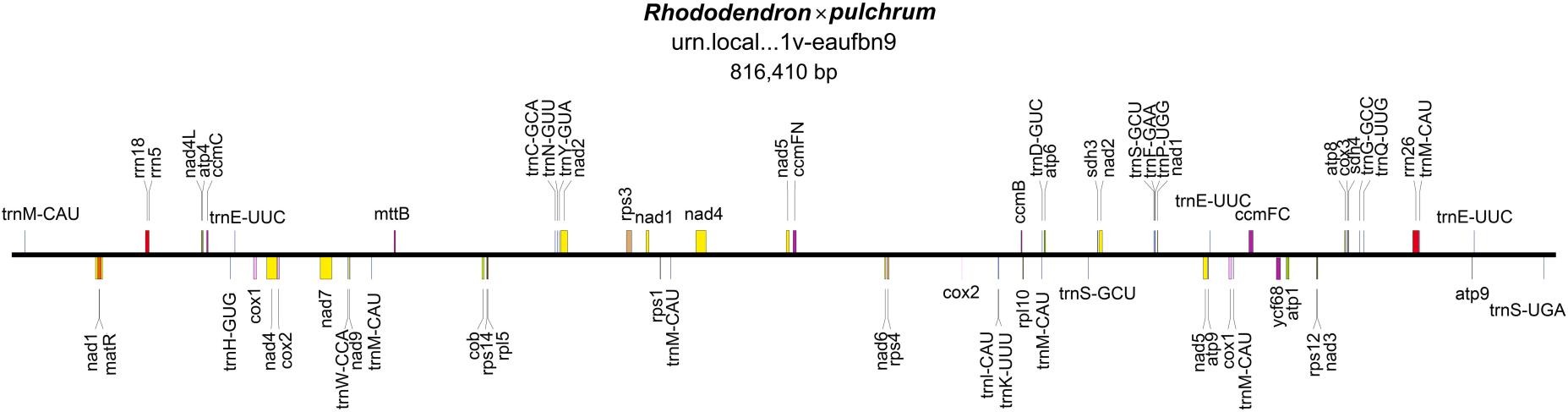


图1

杜鹃花×黄花杜鹃线粒体基因组的基因图谱。

（[**Supplementary Table 7**](#_bookmark14)).其中rrn26基因仅定位一个TR (742，682–742，724 BP)；其他的都位于基因间空间。

## 线粒体基因组比较

杜鹃花科线粒体系统发育树显示，8种植物聚为4组，与它们的分类信息一致。杜鹃花科的5种植物被归为一个分支:西门塔尔牛、山茱萸、蓝萼龙舌兰、低三聚体马齿苋和毛萼龙舌兰。此外，

*R.发现西门塔尔牛是r×普勒姆的姐妹*

（[**Figure 2**](#_bookmark4)).

以r . x pul chrum线粒体基因组为对照，比较了5种杜鹃花科植物的全基因组共线排列，如所示[**Figure 3**](#_bookmark5)。这些相关物种的线粒体基因组在基因区域与r×pul chrum具有高度同源性，其中最保守的基因为atp4、ccmC、nad4、trnC-GCA、trnN- GUU、trnY-GUA、nad2、nad5、trnM-CAU、trnD-GUC和rrn26。

*R.××蚤和西门塔尔牛线粒体基因组较多*

同源覆盖比其他三种杜鹃花科。再者，线粒体基因组组织介于

*R.×普勒姆和西门塔尔牛线粒体基因组显示*

仍然有复杂的变化([**Supplementary**](#_bookmark14)

表2杜鹃花注释基因

叶绿体基因组。

**×**

### 功能基因

rnA，transfer trnA-UGC3，trnC-GCA，trnD-GUC，trnE-UUC，trnF-GAA，trnF-CAU，trnG-GCC，trnG-UCC3，trnH-GUG，trnI-CAU2，trnI-GAU3，trnK-UUU3，trnL-CAA，trnL-UAA3，trnL-UAG，trnM-CAU，trnN-GUU，trnP-UGG，trnQ-UUG，trnR-ACG，trnR-

RNA，核糖体rrn4.5，rrn5，rrn16，rrn23

转录和剪接rpoA，rpoB，rpoC1，rpoC2

小型子单元rps2、rps1、rps4、rps7、rps8、rps11、rps123、rps14、rps15、rps163、rps18、rps191

大亚单位rpl2、rpl14、rpl163、rpl20、rpl22、rpl23、rpl32、rpl33、rpl36

ATP合酶atpA，atpB，atpE，atpF3，atpH，atpI

光系统I psaA、psaB、psaC、psaI、psaJ、ycf1、ycf15、ycf68、pafI4、pafI、pbf1

光系统II psbA、psbB、psbC、psbD、psbE、psbF、psbH、psbI、psbJ、psbK、psbL、psbM、psbT、psbZ

卡尔文循环rbcL

细胞色素复合物petA，petB3，petD3，petG，petL，petN

NADH脱氢酶ndhA、ndhB3、ndhC、ndhD、ndhE、ndhF、ndhG、

*ndhH，ndhI，ndhJ，ndhK*

翻译起始因子infA

Maturase matK

包膜膜蛋白cemA

乙酰辅酶A accD亚单位

### [Figure 3](#_bookmark14)).

中华绒螯蟹线粒体基因组为

c型细胞色素合成ccsA

基因

作为参考，与R. simsii (NC\_053763)比较，使用Mauve和Lastz程序分析线粒体基因组重排和共线性。普勒姆和西门塔尔牛线粒体基因组被划分为63个局部共线区(LCB)。这些lcb在所研究的两种杜鹃的大小和相对位置上有很大差异([**Figure 4**](#_bookmark6)).此外，大多数lcb含有基因序列，有42个基因

1具有2个拷贝的基因；2基因有3个拷贝；3个含有一个内含子的基因；4含有两个内含子的基因。

参与总数([**Supplementary Table 8**](#_bookmark14)).重排分析表明，许多基因在普勒姆和西门塔尔牛线粒体基因组之间发生了重排([**Supplementary Table 8**](#_bookmark14)).

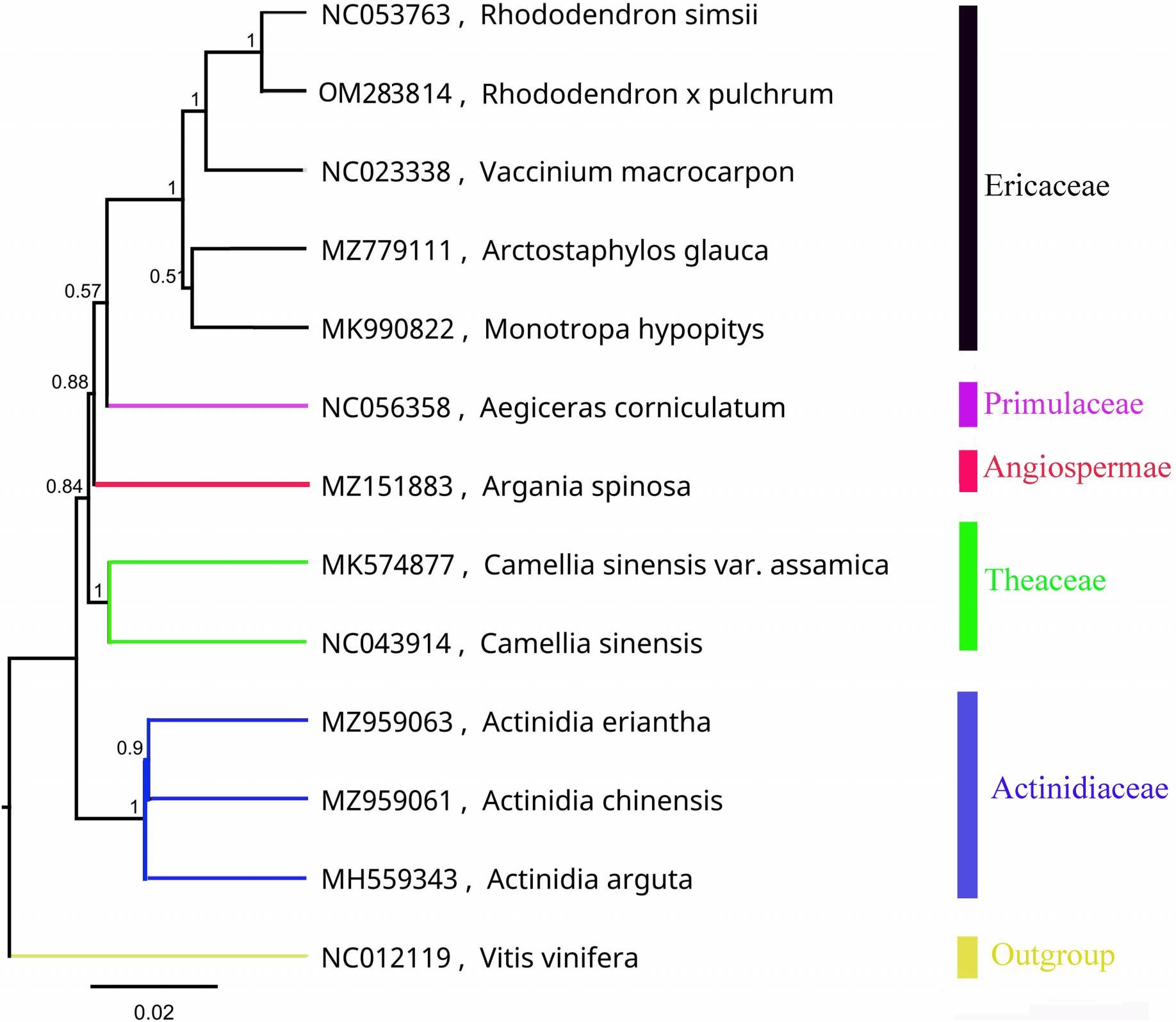


图2

基于线粒体基因组的杜鹃花属物种系统发育树。葡萄(NC012119)被用作外组。

## 叶绿体基因组的组织和特征

结果表明，r×普勒姆叶绿体基因组全长146，941 bp([**Figure 5**](#_bookmark7))并且由于缺乏反向重复序列(IR)，没有采取典型的四方结构的形式。菹草叶绿体基因组的GC含量为35.8%。普勒姆叶绿体全基因组包含119个被注释的基因，即34个tRNA基因、4个rRNA基因和81个蛋白编码基因(mRNA)()[**Table 2**](#_bookmark3)和[**Figure 5**](#_bookmark7)).总共注释了7个蛋白编码基因(ndhB3、petB、petD、atpF、rpl16、rps12和rps16)和6个tRNA (trnA-UGC、trnK-UUU、trnI-GAU、trnG-UCC、trnL-UAA和trnV-UAC)，它们是含内含子的基因([**Table 1**](#_bookmark1)).其中，只有pafI4蛋白编码

基因包含两个内含子，其他的有一个内含子。

基于蛋白编码基因和tRNA基因的密码子使用频率如所示[**Supplementary Table 9**](#_bookmark14)和[**Supplementary Figure 4**](#_bookmark14)。叶绿体基因组由19，278个密码子(65种不同类型)组成

20种氨基酸，其中亮氨酸(Leu)是最常用的氨基酸(10.68%，2，058)，半胱氨酸(Cys)最不丰富(1.12%，215)([**SupplementaryTable 8**](#_bookmark14)).结果表明，r×普勒姆叶绿体基因组偏好以A或U结尾的同义密码子，RSCU > 1([**Supplementary Table 9**](#_bookmark14)和[**Supplementary Figure 4**](#_bookmark14)).

从杜鹃花中鉴定出46个SSR标记

叶绿体基因组，包括41个单核苷酸和5个二核苷酸([**Supplementary Table 10**](#_bookmark14)).包含的所有SSR

图3

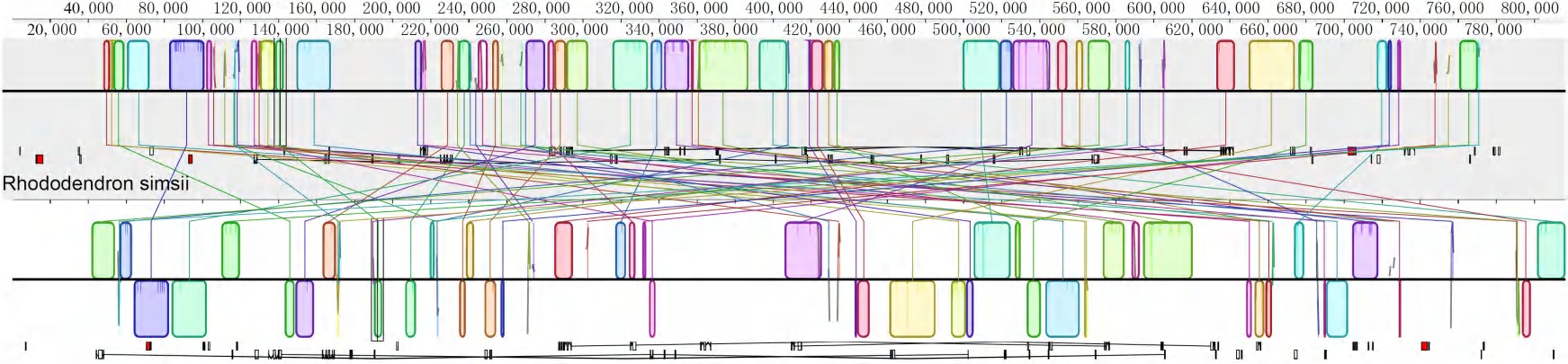
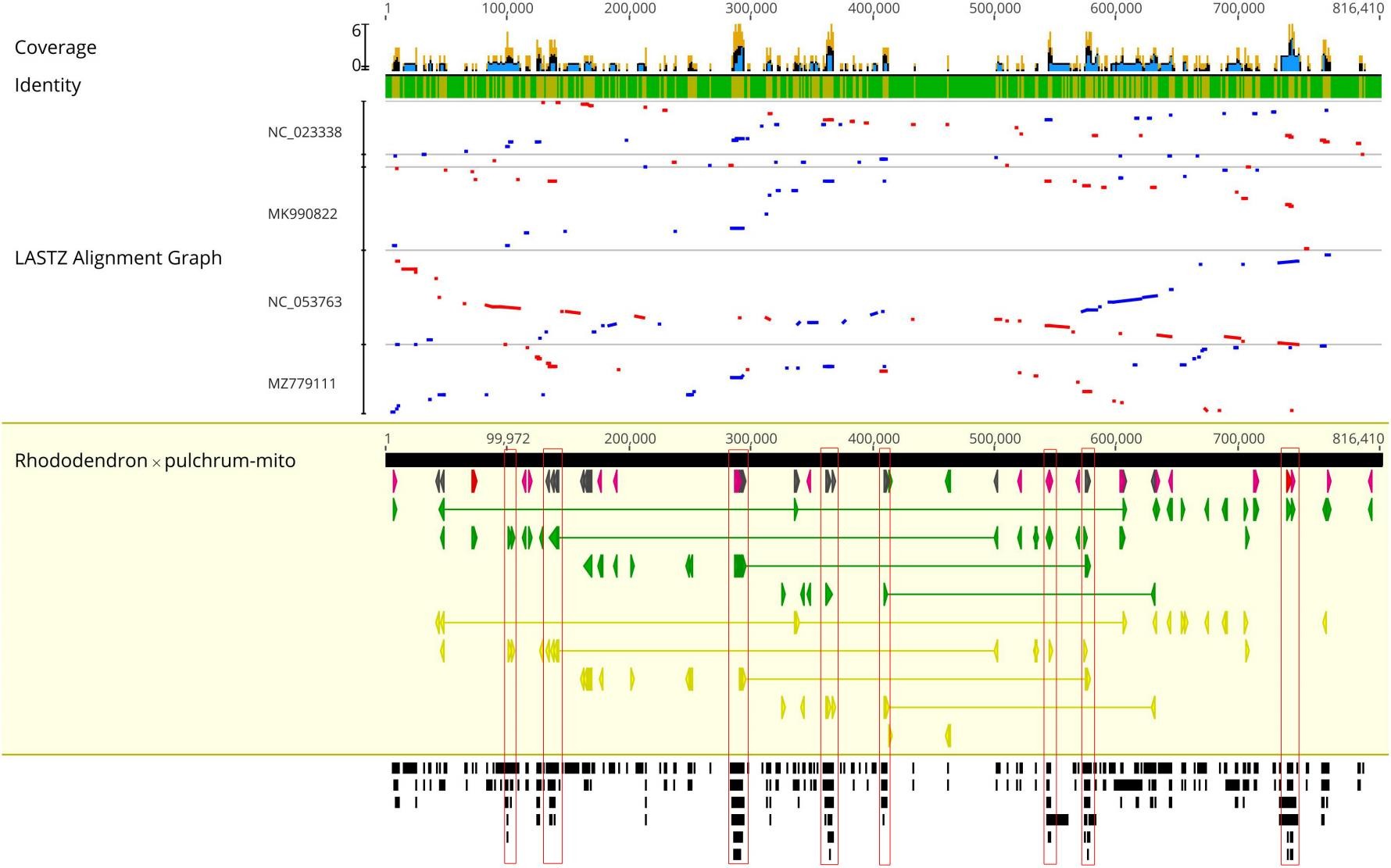
杜鹃花科5种植物线粒体基因组完全共线比对比较。(杜鹃花×黄花杜鹃

作为参考；蓝色条表示同向取向的同源高得分片段对，而红色条表示反向对。黄色箭头表示编码序列，红色箭头表示核糖体RNA基因，紫色箭头表示转移RNA基因，绿色箭头表示蛋白质编码基因，灰色箭头表示外显子区域。黑色方块表示同源性；数量越大，黑色方块的面积越大，并且更密切相关的物种具有同源片段。红色线框是最保守的区域。这些区域在图中从左到右标记，从左到右分布，包含以下基因:atp4和ccmCnad4trnC-GCA、trnN-GUU、trnY-GUA和nad2nad4nad5trnM-CAU和trnD-GUC；nad2rrn26和trnM-CAU)。

图4

杜鹃与西门塔尔牛线粒体全基因组比较的紫红色可视化。[渐进性淡紫色

比对(在Mauve程序中)显示了线粒体基因组之间共享的同源区块，还用线连接了这些区块，指明了同源区块之间的相应位置，以可视化基因排列】。

仅聚腺嘌呤(聚A)或聚乙胺(聚T)，不含鸟嘌呤(G)或胞嘧啶(C)。

## 叶绿体基因组比较

杜鹃花属15种植物的叶绿体基因组进化树显示，这些物种是聚集的

分成三组，这与它们的分类信息是一致的。亚门有五种。处女膜聚为一个分支:R. delavayi var。Delavayi，R. griersonianum，R. henanense亚种。lingbaoense，R. platypodum和R. delavayi。R. molle属于Subgen。五花草被聚集在一个枝上。其他6种属于亚属。铁杉被聚集成一个枝。在这6种植物中，川上草和大田草属于同一类群。

Vireya、R. micranthum和R. concinnum属于同一教派。杜鹃花属、杜鹃花属、杜鹃花属。Tsutsusi。此外，发现ripense是r×pul chrum的姊妹种([**Figure 6**](#_bookmark8)).

## 密码子使用偏向模式与进化

来自杜鹃花科的五个线粒体基因组的密码子使用频率已在[**Supplementary Table 11**](#_bookmark14)，这五个物种中有30∞32个密码子，RSCU > 1，这些密码子以A或T结尾，除了UUG (Leu)以g结尾。密码子GCU (Ala)、CAU (His)和UAA (Ter)是最常见的密码子(RSCU > 1.5)，CAS (His)是这五个物种中最不常见的密码子(RSCU < 0.5)。杜鹃花15个叶绿体基因组的密码子使用频率已在[**Supplementary Table 12**](#_bookmark14)，这15个物种中有30个∞31个密码子，RSCU > 1，这些密码子以A或T结尾，除了UUG (Leu)以g结尾。此外，14个∞19个密码子是最常见的密码子(RSCU > 1.5)，16个∞20个密码子是这15个物种中最不常见的密码子(RSCU < 0.5)。

发育不全分枝杆菌ENC值低于35的基因数，

*动词 （verb的缩写）山茱萸、蓝萼香豆、R. ×普勒姆和R. simsii线粒体基因组为0、0、1 (ccmC)、1 (atp9)和1 (atp9)(*[**Supplementary Table 13**](#_bookmark14)).NEC图分析显示，2009年有10个基因位于或高于标准曲线

5种杜鹃花科植物线粒体基因组，包括atp4、atp6、ccmFc、nad1、nad9、rpl10、rps3、rps10、rps12、rps19，其他基因低于标准曲线线([**Figure 7**](#_bookmark9) 和[**Table 3**](#_bookmark10)).这5个物种中有20个基因都低于标准曲线([**Figure 7**](#_bookmark9)和[**Table 3**](#_bookmark10)).在15种杜鹃叶绿体基因组中发现了3个ENC值低于35的基因，包括petN(发现于15种)、psbI(发现于14种，除了MT239363)、rpl36(发现于11种，除了MT239365、MT239366、NC\_057644和NC\_058233)()[**Supplementary Table 15**](#_bookmark14)).在15种杜鹃花叶绿体基因组中发现的基因atpH、ndhH、petL、psaC、psbM、rpl23和rps11均位于或高于NEC标准曲线([**Figure 8**](#_bookmark11) 和[**Supplementary Table 14**](#_bookmark14)).

## 线粒体和叶绿体基因组间的基因转移

发现线粒体基因组序列(816，410 bp)的长度约是R. × pulchrum叶绿体基因组(146，941 bp)的5.6倍。在长喙库蚊叶绿体和线粒体基因组中鉴定出13个含有基因序列的同源片段([**Table 4**](#_bookmark12)).片段长度在68至1，799 bp之间，保留了> 70%的序列

与它们的原始叶绿体对应物的同一性。这些片段的总长度为4，447 bp，分别占线粒体和叶绿体基因组的0.54%和3.03%。鉴定了三个线粒体基因:trnM- CAU、trnD-GUC和ycf68。共鉴定出13个叶绿体基因，包括trnA-UGC、trnD-GUC、trnI-GAU、trnM-CAU、trnT-UGU、psbF、psbL、psbD、ndhA、ndhF、rrn16、rrn23和ycf68。

# 讨论

## 线粒体基因组结构和大小变异

植物线粒体基因组通常具有典型的环状结构([Sloan et al.](#_bookmark63),[2018](#_bookmark63))，而在某些物种中，线粒体基因组则表现出线性或多染色体结构，如二氧化硅和藜([Backert and Borner](#_bookmark17),[2000](#_bookmark17)；[Sloan et al.](#_bookmark62),[2012](#_bookmark62)；[Sanchez-Puertaet al.](#_bookmark53),[2017](#_bookmark53)).植物线粒体基因组的大小分布广泛，通常在66 ~ 11，000 kb([Wuet al.](#_bookmark77),[2022](#_bookmark77)).经鉴定，中华绒螯蟹线粒体基因组为816，410 bp的谱系DNA分子。西门塔尔牛线粒体基因组显示802，707 bp的谱系DNA分子([Xu et al.](#_bookmark79),[2021](#_bookmark79))，而山茱萸v .显示具有468，115 bp的单个圆形支架([Luis et al.](#_bookmark45),[2019](#_bookmark45)).在植物分类学上，西门塔尔牛和杜鹃花属杜鹃花科，山茱萸属杜鹃花科。11种杜鹃花属植物的集合线粒体基因组大小在425，282至816，410个核苷酸之间；在杜鹃花科的5个物种中，它们的范围从459，678到816，410个核苷酸。先前的研究表明，植物线粒体基因组在物种之间可以表现出很大的差异([Wu et al.](#_bookmark77),[2022](#_bookmark77)).这些结果表明，杜鹃花科物种线粒体基因组的结构和大小的变化差异很大，这可能与物种进化和系统发育的亲缘关系有关。密切相关的物种之间的线粒体基因组的结构和大小更接近，例如

*R.simsii和R. × pulchrum(*[Xu et al.](#_bookmark79),[2021](#_bookmark79))和之间

*C.中华按蚊和中华按蚊变种。assa小鼠(*[Zhang et al.](#_bookmark80),[2019](#_bookmark80))，表明线粒体基因组可能是研究植物进化和鉴定物种分类群的潜在策略([Li et al.](#_bookmark43),[2018](#_bookmark43)；[Zubaer et al.](#_bookmark81),[2018](#_bookmark81)).

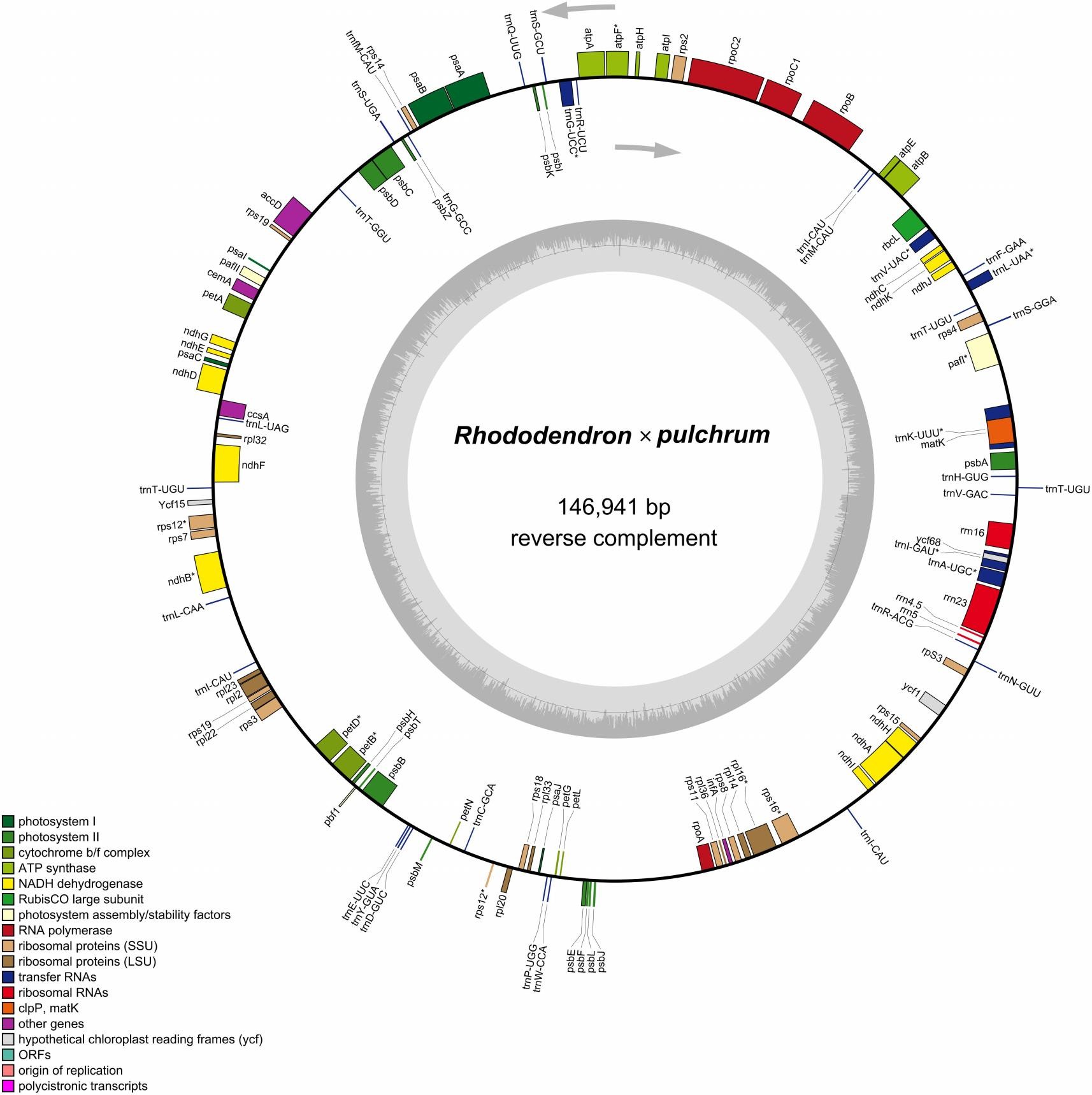
中华绒螯蟹的线粒体基因组最长

到目前为止报道的杜鹃花属物种的基因组，包含36个蛋白质编码基因、3个rRNA基因和24个tRNA基因，类似于西门塔尔牛([Xu et al.](#_bookmark79),[2021](#_bookmark79))和中华按蚊变种。assa小鼠([Zhang et al.](#_bookmark80),[2019](#_bookmark80)).然而，来自杜鹃花科的三种不同物种:杜鹃花、西门塔尔牛和中华杜鹃花的线粒体基因组中的基因的多个拷贝完全不同。assa小鼠。此外，还有更多注释

图5

杜鹃花×黄花杜鹃叶绿体基因组的基因图谱。(位于外圈内外的基因位于

分别为正向和反向。内圈的深灰色条和浅灰色条分别表示G + C和A + T含量)。

中华按蚊变种中的基因。a小鼠线粒体基因组比R. × pulchrum小鼠少，相关基因包括rrn16、rpl2、rpl16、rps7、rps13、rps19和trnfM。已发现植物中的多染色体基因组具有可变的拷贝数([Alverson et al.](#_bookmark15),[2011](#_bookmark15)；[Wu et al.](#_bookmark78),[2015](#_bookmark78))，而线粒体基因拷贝数却没有更详细的数据，这可能是由于迄今为止植物线粒体的数据有限。线粒体基因atp1和matR是保守的，已被用于研究杜鹃花目（Ericales）([Anderberg et al.](#_bookmark16),[2002](#_bookmark16)).这些结果与我们的研究一致。因此，需要对植物线粒体基因组进行进一步研究，这将为这些结构多样的基因组的基因组进化和分子标记提供新的见解。

## 叶绿体基因组结构和大小变异

前期研究已经证实，采用第二代测序技术(Illumina Hiseq Platform)得到的叶绿体全基因组测序数据是可靠的([Zhanget al.](#_bookmark80),[2019](#_bookmark80)；[Xu et al.](#_bookmark79),[2021](#_bookmark79)).使用第二代

测序技术发现，普勒姆叶绿体基因组全长136，249 bp，由于缺失反向重复序列，没有典型的四方结构，GC含量为35.98%。总之，

鉴定出73个功能基因，包括2个rRNA基因、29个tRNA基因和42个蛋白编码基因([Shenet al.](#_bookmark60),[2020](#_bookmark60)).在这项研究中，第二和TGS技术

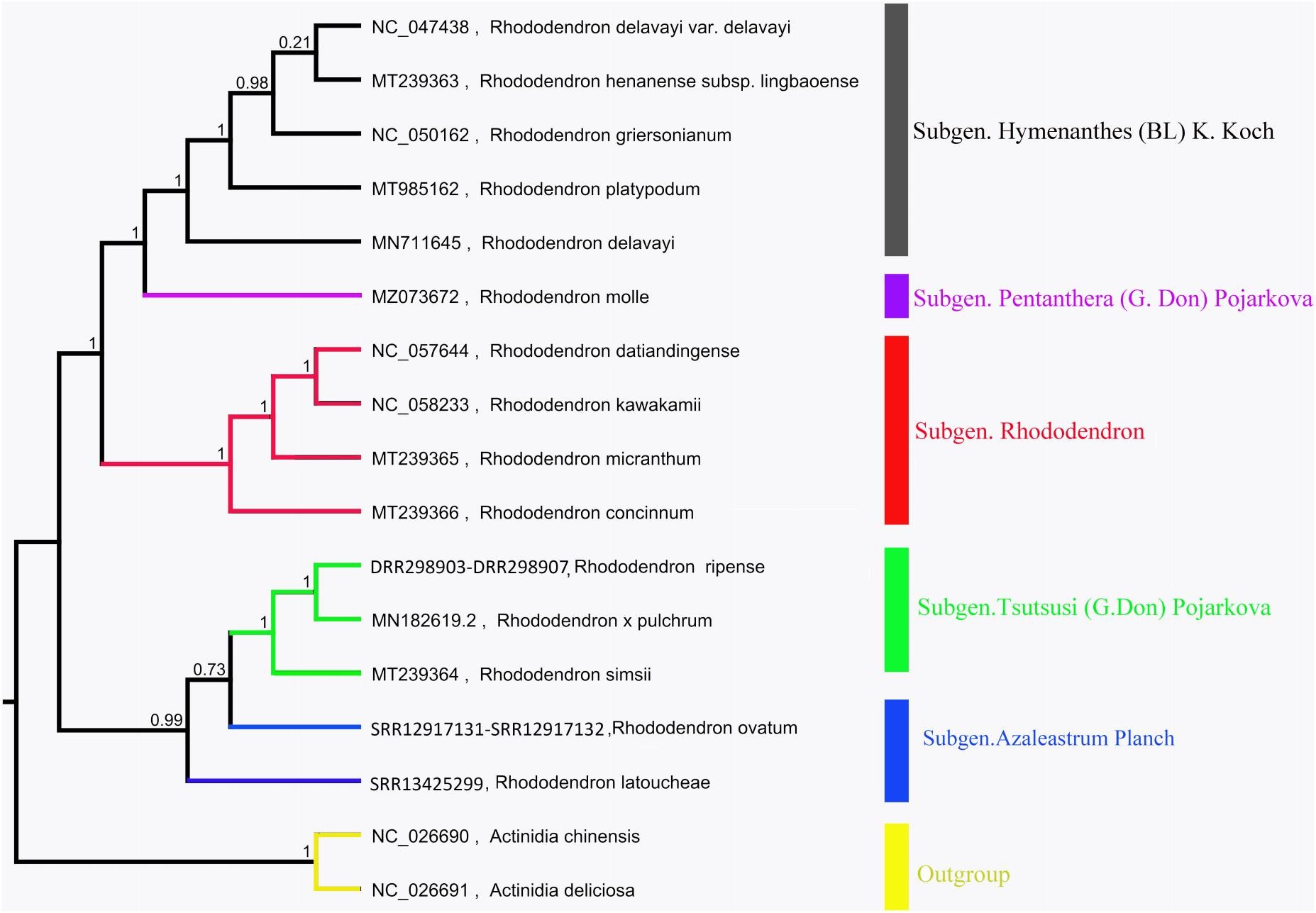


图6

基于叶绿体基因组的杜鹃属物种系统发育树。以美味猕猴桃和中华猕猴桃为外群。

用于重组和重新命名r×pul chrum的完整叶绿体基因组。我们的新结果显示，完整的r×普勒姆叶绿体基因组长度为146，941 bp([**Figure 5**](#_bookmark7));缺乏典型的四方结构，GC含量与以往研究结果相比无明显变化。然而，完整的叶绿体基因组

*R.×花粉含有119个注释基因；比以前的研究多注释了46个基因，包括*

5个tRNA基因、2个rRNA基因和39个蛋白编码基因(mRNA)([**Table 1**](#_bookmark1)).使用Illumina Hiseq平台对之前发表的R. × pulchrum质体基因组进行测序([Shen et al.](#_bookmark59),[2019](#_bookmark59),[2020](#_bookmark60)).本研究采用TGS方法结合第二代测序技术对普氏乳杆菌的完整质体基因组进行了检测。使用Lastz软件比较两种方法的测序结果，以考察结果的准确性。杜鹃×黄花杜鹃的MN182619.1和MN182619.2之间的完整叶绿体基因组共线比对已在中显示[**Figure 9A**](#_bookmark13)。比较结果发现有多个断裂和反转位点，其中6个具有

选择长片段(10，000–30，000 bp)来设计用于PCR实验引物([**Supplementary Table 16**](#_bookmark14)).PCR结果符合预期大小([**Figure 9B**](#_bookmark13))，测序结果与MN182619.2一致([**Figure 9C**](#_bookmark13)).我们的研究表明，TGS技术(牛津纳米孔平台)可以提高先前未组装的基因组区域的覆盖和组装精度，并保证基因组长度和内容的准确性。第三代测序技术可能是解释植物细胞器基因组信息的有用工具([Shearman et al.](#_bookmark58),[2016](#_bookmark58))对于具有长读取长度的TGS方法。

根据已发表的数据，大多数植物的质体基因组通过显示LSC、LSC和两个IRs([Wicke et al.](#_bookmark74),[2011](#_bookmark74)；[Shenet al.](#_bookmark60),[2020](#_bookmark60)).我们的结果表明，R. × pulchrum cp基因组缺少IRs，这与delavayi杜鹃([Li et al.](#_bookmark42),[2020](#_bookmark42)).不同种的质体基因组表现出不同的结构。这说明质体基因组可能是研究进化和鉴定红顶体物种的物种分类群的一种潜在策略([Li et al.](#_bookmark43),[2018](#_bookmark43)；[Zubaer et al.](#_bookmark81),[2018](#_bookmark81)).

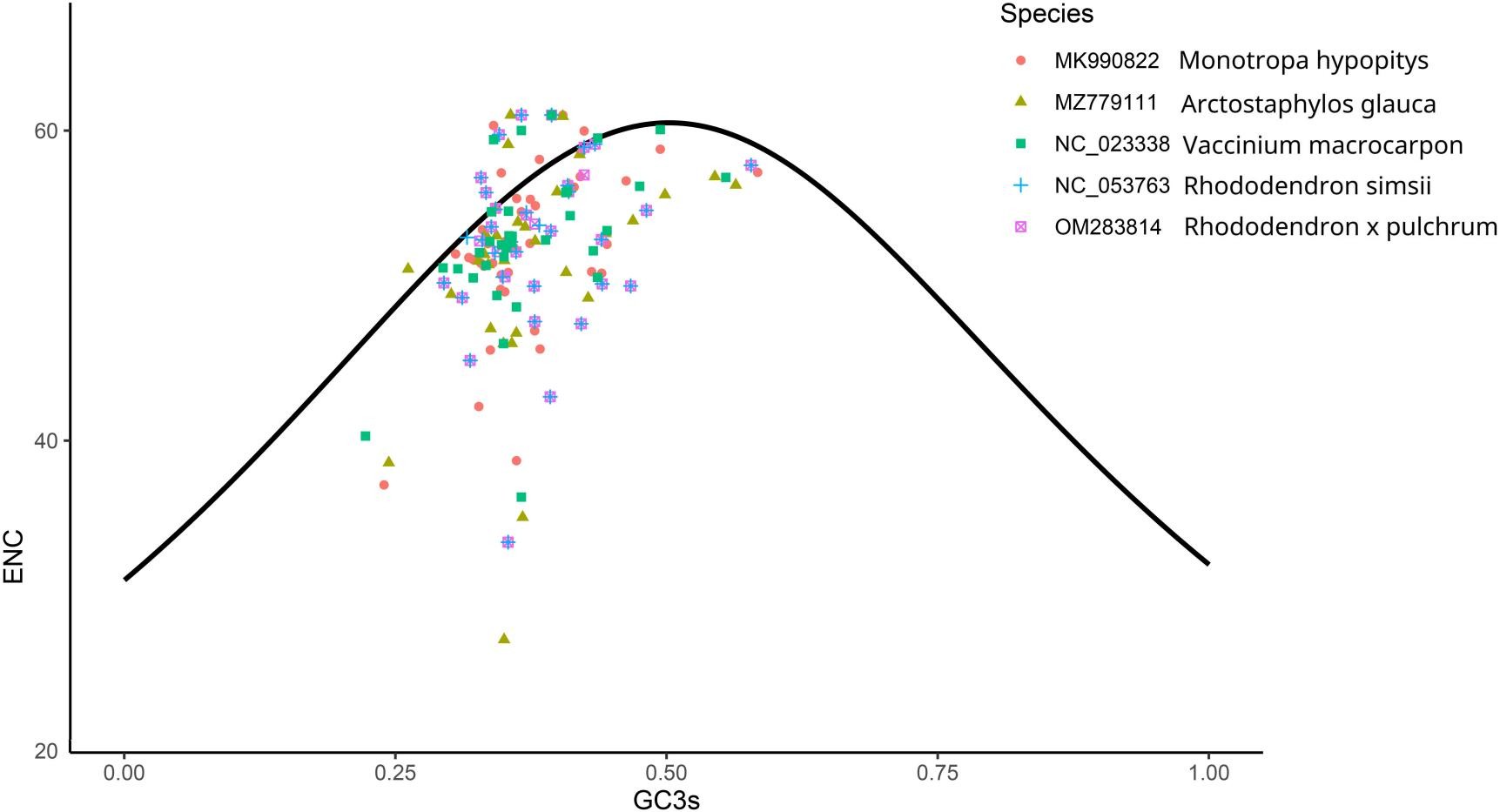


图7

5种杜鹃花属植物线粒体基因组中基因的ENC图分析。(Y轴:基因有效密码子个数值；x轴:基因的GC3含量；标准曲线计算如下:ENC = 2 + GC3 + 29/(GC32 + (1-GC3)2)。

表3 5种杜鹃花科植物线粒体基因组中基因的ENC图。

### 属种基因库

**ENC图分析**

**标准曲线线上或标准曲线线上低于或高于标准曲线线的基因ENC值数**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | **基因编号** | **基因表** |  | **基因编号** | **基因表** |
| *Monotropa* | *Monotropa MK990822*  *发育不全* | 5 | *atp4，nad9，rps10，rps19，rps3* |  | 33 | *atp1、atp6、atp8、atp9、ccmB、ccmC、ccmFc、ccmFn、cob、cox1、cox2、cox3、matR、mttB、nad1、nad2、nad3、nad4、nad4L、nad5、nad6、nad7、rpl2、rpl5、rpl10、rpl16、rps1、rps4、rps12、rps13、rps19、sdh3、sdh4* |
| *Arctostaphylos* | *Arctostaphylos MZ779111*  *格劳卡* | 四 | *atp4，nad1，nad9，rpl10* |  | 29 | *atp1、atp8、atp9、ccmB、ccmC、ccmFc、ccmFn、cox1、cox2、cox3、cytB、matR、mttB、nad2、nad3、nad4、nad5、nad6、nad7、rpl2、rpl16、rps1、rps4、rps10、rps12、rps13、rps15、rps19、sdh4* |
| *杜鹃花* | *杜鹃× OM283814*  *pulchrum* | 5 | *atp4、atp6、nad9、rpl10、rps12* |  | 28 | *atp1、atp8、atp9、ccmB、ccmC、ccmFc、ccmFn、cob、cox1、cox2、cox3、matR、mttB、nad1、nad2、nad3、nad4、nad4L、nad5、nad6、nad7、* |
|  | *杜鹃NC\_053763*  *映山红* | 5 | *atp4、atp6、nad9、rpl10、rps12* |  | 28 | *rpl5、rps1、rps3、rps4、rps14、sdh3、sdh4*  *atp1、atp8、atp9、ccmB、ccmC、ccmFc、ccmFn、cob、cox1、cox2、cox3、matR、mttB、nad1、nad2、nad3、nad4、nad4L、nad5、nad6、nad7、* |
| *越橘科* | *痘苗NC\_023338*  *山茱萸* | 四 | *atp4，ccmFc，nad9，rpl10* |  | 29 | *rpl5、rps1、rps3、rps4、rps14、sdh3、sdh4*  *atp1、atp8、atp9、ccmB、ccmC、ccmFn、cob、cox1、cox2、cox3、matR、mttB、nad1、nad2、nad3、nad4、nad5、nad6、nad7、rpl2、rpl16、* |
|  |  |  |  |  |  | *rps1、rps4、rps10、rps12、rps13、rps15、rps19、sdh4* |

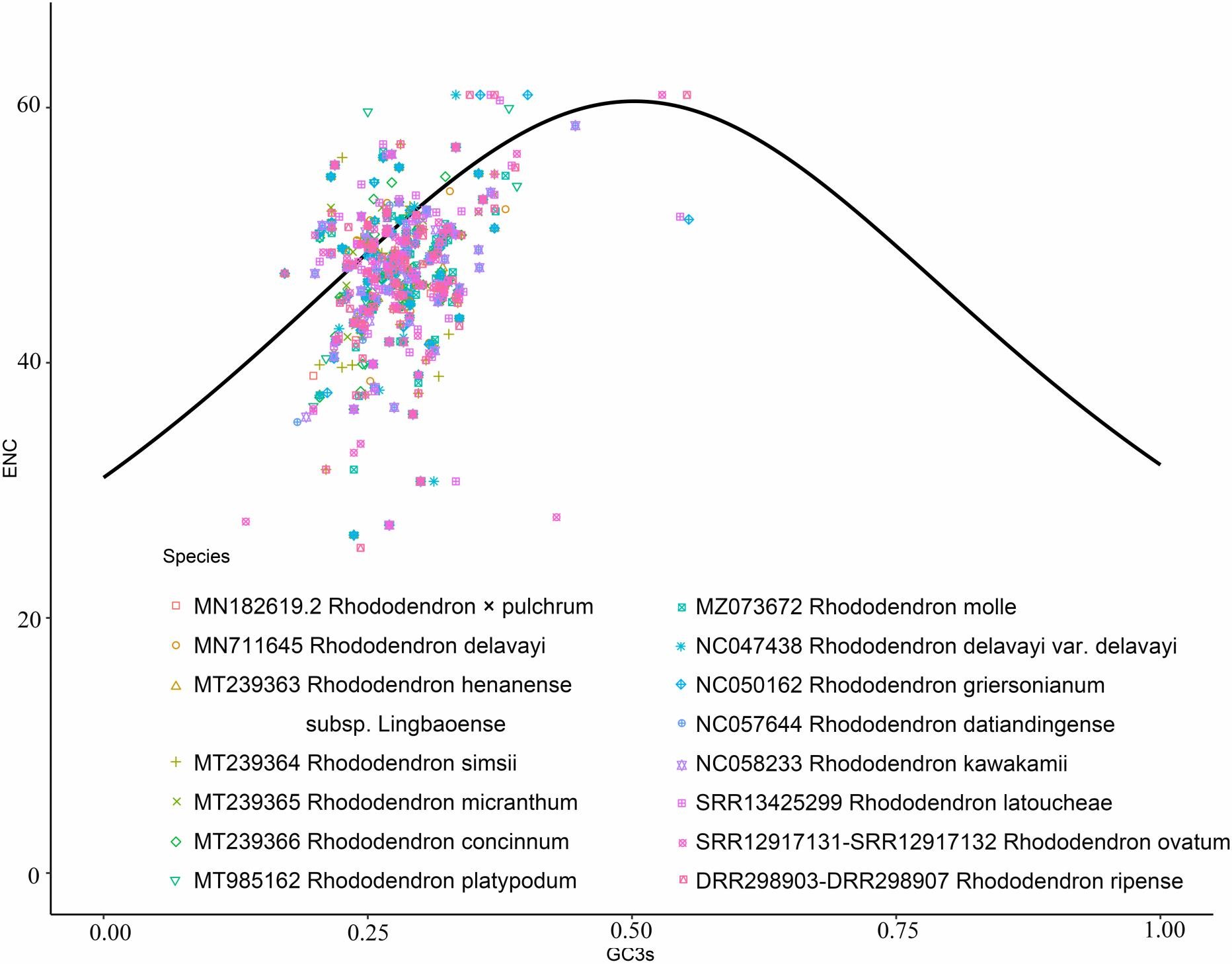


图8

12种杜鹃叶绿体基因组中基因的ENC图分析。(Y轴:基因有效密码子个数值；x轴:基因的GC3含量；标准曲线计算如下:ENC = 2 + GC3 + 29/(GC32 + (1-GC3)2)。

## 系统发育和线粒体基因组比较

8种杜鹃花属植物和15种杜鹃花属植物的线粒体和叶绿体系统发育树的结果与其分类信息一致。这些结果说明细胞器基因组序列数据适用于研究植物的系统发育([Wu andGe](#_bookmark76),[2012](#_bookmark76)；[Thomas et al.](#_bookmark68),[2017](#_bookmark68)；[Shen et al.](#_bookmark60),[2020](#_bookmark60)).本研究中的系统发育树表明，ripense是r××pul chrum的姊妹种([**Figures 2**](#_bookmark4),[**6**](#_bookmark8)).以往研究

表明R. ×普勒姆栽培品种拥有R. ripense的cpDNA

（[Shirasawa et al.](#_bookmark61),[2021](#_bookmark61))中发现，并且认为ripense r .被认为是R. × pulchrum的推定祖先物种之一([Scariotet al.](#_bookmark55),[2007](#_bookmark55)；[Meanchaipiboon et al.](#_bookmark46),[2021](#_bookmark46))，这些发现与我们的研究一致。以锦绣杜鹃线粒体基因组为对照；在来自杜鹃花科的五个物种之间比较的全基因组共线比对发现与杜鹃花的高度同源性

在基因区与间隔区相比。这说明在杜鹃花科物种的线粒体基因组中，基因区域比间隔区域更保守。本研究中鉴定的11个保守基因(atp4，ccmC，nad4，trnC- GCA，trnN-GUU，trnY-GUA，nad2，nad5，trnM-CAU，trnD- GUC和rrn26)将有助于理解杜鹃花科植物的进化机制和物种分类群([Li et al.](#_bookmark43),[2018](#_bookmark43)；[Zubaer et al.](#_bookmark81),[2018](#_bookmark81)).

中华鼢鼠和西门塔尔牛的线粒体基因组

被鉴定为63种LCBs这些液晶显示器在尺寸和相对位置上有很大不同([**Figure 4**](#_bookmark6)).此外，大多数lcb含有基因序列。我们的结果表明，基因大小没有显著差异([Xu et al.](#_bookmark79),[2021](#_bookmark79)).尽管许多基因在普勒姆和西门塔尔牛线粒体基因组之间进行了重排，但它们的线粒体基因组存在非常复杂的变异，这可能导致杂交和基因组加倍事件。这些结果与之前显示高变异的研究一致

表4杜鹃花线粒体和叶绿体基因组间的基因转移。

|  |  |  |  |
| --- | --- | --- | --- |
| **序列名称** | **线粒体基因组**  **基因序列位置(bp)** | **叶绿体基因组**  **基因序列位置(bp)** | **身份(%)** |
| S0 | *trnM-CAU 6，870–6，937* | *trnT-UGU 104，61–10，530* | 72.86 |
| S1 | *trnM-CAU 544，505–544，570* | *trnM-CAU 20，294–20，361* | 89.71 |
| S2 | *trnD-GUC 98，655–98，779* | *trnD-GUC 97，859–97，983* | 86.4 |
| S3 | / 580,657–580,726 | *psbF 108，602–108，671* | 90 |
| 第四心音 | / 580,735–580,862 | *psbFpsbL 108，685–108，812* | 90.63 |
| 表面抗原-5 | *ycf68 669，585–671，383* | *ycf68trnA-UGC(外显子1)；140,647–142,445*  *trnI-GAU(外显子2)；trnI-GAU*  (内含子1)；trnA-UGC(内含子1) | 100 |
| S6 | / 57,645–75,825 | *psbD 52，426–53，035* | 99.34 |
| 正常人血清中的一种蛋白质成分 | / 478,324–478,417 | *ndhF 73，473–73，566* | 85.11 |
| S8 | / 439,289–439,591 | *ndhA(内含子1) 127，789–128，091* | 99.67 |
| S9 | / 438,679–439,302 | *ndhAndhA(内含子1) 128，097–128，720* | 99.68 |
| S10 | / 341,227–341,397 | *rrn23 139，234–139，404* | 95.91 |
| S11 | / 71,461–71,598 | *rrn16 143，386–143，523* | 89.13 |
| S12 | / 70,854–71,050 | *rrn16 143，938–144，134* | 78.17 |

和植物线粒体基因组中的重组([Galtier](#_bookmark27),[2011](#_bookmark27)).此外，植物线粒体基因组中的基因顺序是高度可变的([Fischer et al.](#_bookmark25),[2006](#_bookmark25)；[Schoch et al.](#_bookmark56),[2015](#_bookmark56)；[Li et al.](#_bookmark44),[2017](#_bookmark44)).线粒体中基因的排列可用于评估不同物种之间的系统发育关系([Liet al.](#_bookmark44),[2017](#_bookmark44)).

## 密码子使用偏向模式与进化

密码子使用偏性通常用于动物和昆虫分析不同物种的系统发育和进化，然而关于植物([Sun et al.](#_bookmark66),[2009](#_bookmark66)；[Huang and Ma](#_bookmark32),[2018](#_bookmark32)).ENC值低(ENC < 35)的线粒体叶绿体基因的基因密码子使用偏向；[**Supplementary Tables 13, 14**](#_bookmark14))可能是由突变引起的，而其他基因的密码子使用偏向是由选择或其他因素引起的([Wright](#_bookmark75),[1990](#_bookmark75)).因此，我们只能推断出10个杜鹃花科线粒体基因(atp4、atp6、ccmFc、nad1、nad9、rpl10、rps3、rps10、rps12和rps19)和7个杜鹃花叶绿体基因(atpH、ndhH、petL、psaC、psbM、rpl23和rps11)的基因密码子使用偏向受突变影响，而其他基因的密码子使用偏向是经过自然或人工选择的。线粒体基因NEC图谱分析表明，普勒姆杜鹃和西门塔尔牛之间的基因密码子使用是一致的，叶绿体基因组中同一亚属或亚属种的杜鹃花的基因密码子使用也是一致的。这些结果表明，线粒体和叶绿体基因组的NEC图谱分析可能适用于杜鹃花科植物的进化分析。

## 线粒体和叶绿体基因组间的基因转移

先前的研究已经检测到不同基因组之间的细胞内基因转移，这已经通过对核基因组、线粒体基因组和叶绿体基因组的测序分析([Timmis et al.](#_bookmark70),[2004](#_bookmark70)；[Nguyenet al.](#_bookmark47),[2020](#_bookmark47)).这些研究大多集中在被子植物细胞器中细胞核DNA的基因转移([Smith](#_bookmark64),[2011](#_bookmark64)；[Park et al.](#_bookmark48),[2014](#_bookmark48)).随着新技术的发展，越来越多的细胞器基因组数据被发表和分析，叶绿体到线粒体的基因转移被认为是长期进化的特征([Gui et al.](#_bookmark30),[2016](#_bookmark30)；[Nguyen et al.](#_bookmark47),[2020](#_bookmark47)).在本研究中，在的叶绿体和线粒体基因组之间鉴定了13个含有基因序列的同源片段

*R.× pulchrum(*[**Table 4**](#_bookmark12))，这可能是

细胞器基因组间的水平基因转移。在R. × pulchrum线粒体基因组中，10个叶绿体基因未被注释，包括3个tRNA (trnA- UGC、trnI-GAU和trnT-UGU)、2个核糖体基因(rrn16和rrn23)、3个光合基因(psbF、psbL和psbD)和2个NADH脱氢酶基因(ndhA和ndhF)，但包含它们的同源片段([**Table 4**](#_bookmark12)).核糖体基因参与核糖体复合体的合成，核糖体复合体参与正常细胞功能所需的蛋白质([Depamphilis and Palmer](#_bookmark23),[1990](#_bookmark23)；[Wang et al.](#_bookmark72),

[2019](#_bookmark72)).光合基因和NADH脱氢酶基因参与光合系统ⅱ和叶绿体NADH脱氢酶复合体的合成，参与光合作用([Sazanov et al.](#_bookmark54),[1998](#_bookmark54)；

图9

鉴定了MN182619.1和MN182619.2之间的结果。(A)比较了之间的完整叶绿体基因组共线比对

杜鹃花的MN182619.1和MN182619.2。MN182619.1作为参考；蓝色条表示同源的高得分

在同向方向上的线段对，而红色条表示反向的线段对。黄色箭头表示编码序列，红色箭头表示核糖体RNA基因，紫色箭头表示转移RNA基因，绿色箭头表示蛋白质编码基因，灰色箭头表示外显子区域。黑色方块表示同源性；数字越大，黑色方块的面积越大，越多

56，637，63，979，115，230，125，853和145，493 bp位置的MN182619.2. (B) PCR产物检测。(位于21，232、

密切相关的物种有同源片段。红色线框是最保守的区域。位于21，232，

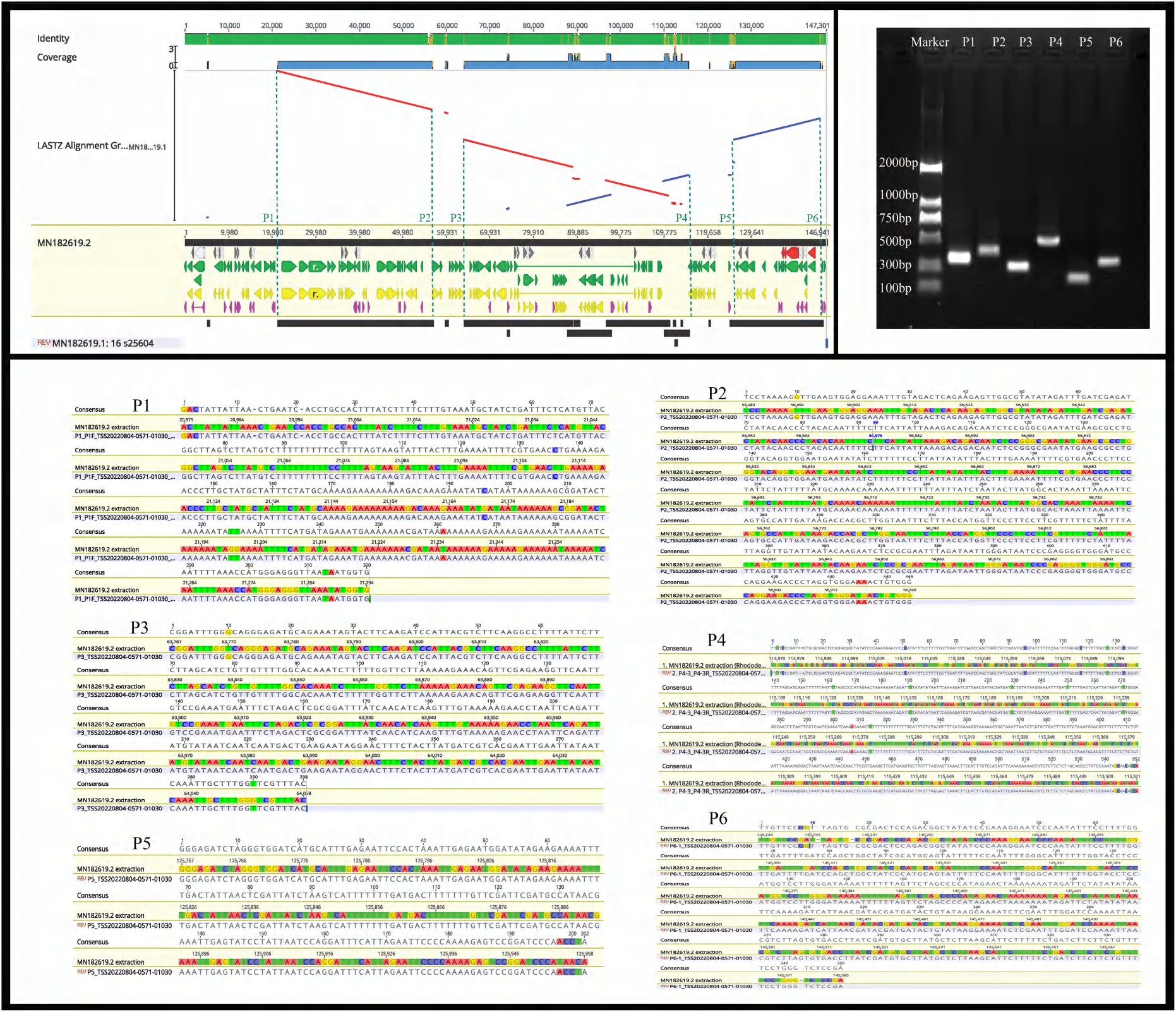
MN182619.2的56，637，63，979，115，230，125，853和145，493 bp位置；P1 P6的预期规模是355，444，300，581，201和

348 bp)。(三)与第三代联合组装的质体基因组fasta (MN182619.2)比较的PCR结果

使用第二代测序方法。[a∞E(P1∞P6):从共线对齐的结果中随机选择6个断裂位点

MN182619.1和MN182619.2的(LASTZ)，如所示[Figure 5](#_bookmark7)。上面的数字表示PCR结果，下面的数字表示用第三代和第二代测序方法组合的质体基因组fasta点]。

[Cai et al.](#_bookmark20),[2021](#_bookmark20)).推测这些与光合作用相关的基因，如psbF、psbL、psbD、ndhA、ndhF等，可能是叶绿体基因转移到r×pul chrum线粒体的结果。需要对植物线粒体和叶绿体基因组进行进一步研究，这将为这些结构多样的基因组的基因组进化、系统发育关系和分子标记提供新的见解。



# 结论

本研究检测了锦绣杜鹃线粒体全基因组，并重组了叶绿体基因组。基因组组织特征，基因组与相关

研究了线粒体和叶绿体基因组之间的物种和基因转移。比较系统发育关系、密码子使用偏误模式和进化与相关

物种。R. × pulchrum结果表明，杜鹃花科植物细胞器基因组变异非常复杂。我们的结果将有助于理解杜鹃花科植物的进化机制和物种分类群的鉴定。

# 数据可用性声明

本研究中展示的数据集可在在线资料库中找到。存储库的名称和注册号可在文章/[**Supplementarymaterial**](#_bookmark14)。

# 作者贡献

JS和SJ:概念化、验证、资源、数据管理和写作——初稿准备、审查和编辑。XL和ML:方法学。ML、XH和HC:软件。SJ:融资收购。所有作者都已阅读并同意手稿的出版版本。

# 提供资金

本研究获得国家重点研发计划项目(2019YFE0118900)、国家自然科学基金项目(31971641和32201608)、浙江省自然科学基金项目资助

(LY16C160011)，浙江A&F大学济阳学院授予(RQ1911B07)。

# 利益冲突

作者ML受雇于生物与数据生物技术有限公司。

其余作者声明，研究是在不存在任何可能被视为潜在利益冲突的商业或金融关系的情况下进行的。

# 出版商声明

本文中表达的所有权利主张仅是作者的权利主张，不一定代表其附属机构的权利主张，也不代表出版商、编辑和评论者的权利主张。本文中可能评估的任何产品，或制造商可能提出的任何索赔，都没有得到出版商的保证或认可。

# 补充材料

本文的补充材料可在以下网站上找到:<https://www.frontiersin.org/articles/10.3389/fpls.2022.969765/full#supplementary-material>

# 参考

a . j . alverson，Rice，D. W .、Dickinson，s .、Barry，k .和Palmer，J. D. (2011)。黄瓜细菌大小的多染色体线粒体基因组的起源和重组。植物细胞23，2499–2513。[doi: 10.1105/tpc。111.087189](https://doi.org/10.1105/tpc.111.087189)

Anderberg，A. A .、Rydin，c .、llersjö，M. (2002年)。杜鹃花目系统发育关系:质体和线粒体基因组中5个基因的分子数据分析。我……J. Bot。89, 677–687.[doi: 10.3732/ajb.89。4.677](https://doi.org/10.3732/ajb.89.4.677)

Backert和t . Borner(2000年)。高等植物藜线粒体中DNA复制和重组的噬菌体T4样中间体。柯。Genet。37, 304–314.[doi:10 . 1007/s0029400505053](https://doi.org/10.1007/s002940050532)

Bankevich，a .，Nurk，s .，Antipov，d .，Gurevich，A. A .，Dvorfin，m .，库利科夫，

A.s .，等人(2012年)。黑桃:一种新的基因组组装算法及其在单细胞测序中的应用。J. Comput。Biol。19, 455–477.[doi: 10.1089/cmb.2012。0021](https://doi.org/10.1089/cmb.2012.0021)

Benson，G. (1999年)。串联重复序列探测仪:一种分析DNA序列的程序。核酸第27，573-580号决议。[doi: 10.1093/nar/27.2.573](https://doi.org/10.1093/nar/27.2.573)

蔡，X. L .、Landis、J. B .、王、H. X .、王、J. H .、朱、Z. X .、王、

H.F. (2021年)。苯乙烯科(杜鹃花科)的质体结构和系统发育关系。BMC Ecol。Evol。21:103.[doi: 10.1186/s12862-021-01827-4](https://doi.org/10.1186/s12862-021-01827-4)

陈，s .，周，y .，陈，y .，顾，J. (2018)。fastp:一种超快速的多功能fastk预处理器。生物信息学34，i884–i890。[doi: 10.1101/274100](https://doi.org/10.1101/274100)

Darling，A. C. E .，Mau，b .，Blattner，F. R .，和Perna，N. T. (2004)。淡紫色:保守基因组序列与重排的多重比对。基因组决议14，1394-1403。[doi: 10.1101/gr.2289704](https://doi.org/10.1101/gr.2289704)

戴潘菲利斯，C. W .，帕尔默，J. D. (1990)。寄生开花植物质体基因组中光合和氯呼吸基因的缺失。Nature 348，337–339。[doi: 10.1038/348337a0](https://doi.org/10.1038/348337a0)

Doyle，J. J. (1987年)。一种用于少量新鲜叶片组织的快速DNA分离程序。植物化学。公牛。19, 11–15.

Fischer，g .，Rocha，E. P .，Brunet，f .，Vergassola，m .，Dujon，B. (2006)。半子囊菌酵母谱系之间基因组重排的高度可变率。PLoS Genet。2:e32。[doi:10.1371/journal . pgen . 0020032](https://doi.org/10.1371/journal.pgen.0020032)

加勒，1985年。杜鹃花。波特兰，或:木材出版社，1985。

Galtier，N. (2011)。植物线粒体DNA有趣的进化动力学。BMC Biol。9:61.[doi: 10.1186/1741-7007-9-61](https://doi.org/10.1186/1741-7007-9-61)

Gray，M. W. (2012)。线粒体进化。CSH透视。Biol。4:a011403。[doi:10.1126/science . 283 . 5407 . 1476](https://doi.org/10.1126/science.283.5407.1476)

Greiner，s .，Lehwark，p .，Bock，R. (2019)。细胞器基因组绘图(OGDRAW)1.3版。1:细胞器基因组图形可视化扩展工具包。核酸Res. 47，W59–W64。[doi: 10.1093/nar/gkz238](https://doi.org/10.1093/nar/gkz238)

桂，S. T .，吴，Z. H .，张，H. Y .，郑，Y. Z .，朱，Z. X .，梁，D. Q .

等人(2016年)。莲的线粒体基因组图谱揭示了古老的进化特征。Sci。代表6:11。[doi: 10.1038/srep30158](https://doi.org/10.1038/srep30158)

Harris，R. S. (2007)。改进的基因组DNA成对排列网络。博士论文。宾夕法尼亚州州立学院:宾夕法尼亚州州立大学。

黄，j .，马，T. (2018)。两种肉蝇线粒体基因组的系统发育比较分析

和麻蝇科的进化时间尺度。里面的J. Biol。Macromol。120, 1955– 1964.[doi:10 . 1016/j . ijbiomac . 2018 . 10 . 001](https://doi.org/10.1016/j.ijbiomac.2018.10.001)

黄，j，梁，X. M .，萱，Y. K .，耿，C. Y .，李，Y. X. (2017)。bgi eq-500测序仪的参考人类基因组数据集。Gigascience 6，1–9。[doi:10.1093/gig science/gix 024](https://doi.org/10.1093/gigascience/gix024)

Jansen，R. K .，Saski，c .，Lee，s .，Hansen，A. K .，Danielle h .(2011)。三种蔷薇属植物(栗、李和可可)的完整质体基因组序列:rpl22至少两次独立转移至细胞核的证据。摩尔。Biol。Evol。28, 835–847.[doi: 10.1093/molbev/msq261](https://doi.org/10.1093/molbev/msq261)

Katoh和d . m . Stanley(2013年)。MAFFT多序列比对软件版本7:性能和可用性改进。摩尔。Biol。Evol。30, 772–780.[doi: 10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010)

kea ls，m .，Moir，r .，Wilson，a .，Stones-Havas，s .，Cheung，m .，Sturrock，s .，et al. (2012)。Geneious basic:一个用于组织和分析序列数据的集成和可扩展的桌面软件平台。生物信息学28，1647-1649。[doi:10.1093/生物信息学/bts199](https://doi.org/10.1093/bioinformatics/bts199)

Knoop，V. (2004年)。陆地植物的线粒体DNA:系统发育视角的特殊性。柯。Genet。46, 123–139.[doi: 10.1007/s00294-004-](https://doi.org/10.1007/s00294-004-0522-8)

0522-[8](https://doi.org/10.1007/s00294-004-0522-8)

Kobayashi，n .，Nakatsuka，a .，Ohta，h .，Kurashige，y .，Huylenbroeck，J. V .，Scariot，v .，et al. (2021)。三叶杜鹃叶绿体基因组对常绿杜鹃花品种发育的贡献。霍特。J. 90，223-231。[doi: 10.2503/hortj。UTD-251](https://doi.org/10.2503/hortj.UTD-251)

Koren，s .，Walenz，B. P .，Berlin，k .，Miller，J. R .，Bergman，N. H .，和Phillippy，

A.M. (2017年)。Canu:通过自适应k-mer加权和重复分离，可扩展且精确的长读汇编。基因组决议27，722-736。[doi: 10.1101/gr。215087.116](https://doi.org/10.1101/gr.215087.116)

Langmead，b .，和Salzberg，S. L. (2012)。与领结2快速间隙读取对齐。纳特。方法9，357–359。[doi: 10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923)

李，H. (2018)。最小值2:核苷酸序列的成对比对。生物信息学34，3094-3100。[doi:10.1093/生物信息学/bty191](https://doi.org/10.1093/bioinformatics/bty191)

李，h，郭，q，李，q，杨，L. (2020)。长序列分析表明，德令哈杜鹃质体基因组含有广泛的重复序列，光合杜鹃花科植物质体基因组之间存在重组。PeerJ 8:e9048。[doi:10.7717/peerj.9048](https://doi.org/10.7717/peerj.9048)

李，q，杨，m，陈，c，熊，c，金，x，浦，Z. Q，等(2018)。药用真菌硫磺菌线粒体全基因组的特征及系统发育分析。Sci。代表8:9104。[doi: 10.1038/](https://doi.org/10.1038/s41598-018-27489-9)s 14598-[018-27489-9](https://doi.org/10.1038/s41598-018-27489-9)

李，x .，徐，X. X .，张，f .，王，m .，徐，y .，唐，d .，等(2017)。线粒体靶向抗氧化剂mitoq通过nrf2/pink1改善了糖尿病肾病中有丝分裂介导的肾小管损伤。氧化还原生物。11, 297–311.[doi:10.1016/j . redox . 2016 . 12 . 022](https://doi.org/10.1016/j.redox.2016.12.022)

Luis，D. G .，Lorraine，R. B .，Jessica，r .，Tyler，s .，和Juan，Z. (2019)。Pacbio测序显示美国蔓越橘(山茱萸越橘)的细胞器基因组完全相同。)和一个野生亲戚。基因10:291。[doi: 10.3390/](https://doi.org/10.3390/genes10040291)基因[s10040291](https://doi.org/10.3390/genes10040291)

Meanchaipiboon，s .，Kobayashi，n .，和Nakatsuka，A. (2021)。从类黄酮3 '，5 '羟化酶基因序列推断的平度杜鹃花品种及其推定亲本之间的遗传关系。霍特。J. 90，114-121。[doi: 10.](https://doi.org/10.2503/hortj.UTD-228)

2503/hort[j.UTD-228](https://doi.org/10.2503/hortj.UTD-228)

Nguyen，V. B .，Giang，V. N. L .，Waminal，N. E .，Park，H. S .，Kim，N. H .，Jang，w .，et al. (2020)。七种人参叶绿体基因组的综合比较分析及基于种特异性单核苷酸多态性标记的鉴定系统的开发。j .人参第44，135-144号决议。[doi: 10.1016/j.jgr.2018.06.003](https://doi.org/10.1016/j.jgr.2018.06.003)

Park，s .，Ruhlman，T. A .，Sabir，J. S. M .，Mutwakil，M. H. Z .，Baeshen，

米（meter的缩写））n .，Sabir，M. J .，et al. (2014)。药用植物狭叶鼠李(夹竹桃科)的细胞器基因组的完整序列，以及不同星状体之间线粒体基因组进化的对比模式。BMC Genom。15:405.[doi: 10.1186/1471-2164-15-405](https://doi.org/10.1186/1471-2164-15-405)

帕金森，C. L .，割草机，J. P .，邱，Y. L .，Shirk，A. J .，Song，k .，Young，N. D .，等(2005).植物老鹳科线粒体取代率的多次重大增减。BMC Evol。Biol。5:73.[doi: 10.1186/1471-2148-5-73](https://doi.org/10.1186/1471-2148-5-73)

Peden和John，F. (2000年)。密码子使用分析，第9卷。诺丁汉:诺丁汉大学，73-74。[doi: 10.0.3.238/expr.1997.4185](https://doi.org/10.0.3.238/expr.1997.4185)

Price，M. N .，Dehal，P. S .，Arkin，A. P. (2010)。fast tree 2–适用于大型路线的近似最大似然树。PLoS One 5:e9490。[doi: 10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490)

Rice，p .，Longden，I .，和Bleasby，A. (2000)。浮雕:欧洲分子生物学开放软件套件。趋势基因。16, 276–277.[doi: 10.1016/S0168-9525(00)02024-2](https://doi.org/10.1016/S0168-9525(00)02024-2)

Sanchez-Puerta，M. V .，García，L. E .，Wohlfeiler，j .和Ceriotti，L. F. (2017)。全寄生植物中天然线粒体基因被外来同源物无与伦比的取代。新植物醇。214, 376–387.[doi: 10.1111/nph.14361](https://doi.org/10.1111/nph.14361)

Sazanov，L. A .、Burrows，P. A .、Nixon，P. J. (1998)。质体ndh基因编码一种NADH特异性脱氢酶:从豌豆类囊体膜中分离出一种复合物I类似物。继续。纳特。阿卡德。Sci。U.S.A. 95，1319–1324。[doi:10.1073/pnas.95.3.1319](https://doi.org/10.1073/pnas.95.3.1319)

Scariot，v .，Handa，t .，和Riek，J. D. (2007年)。利用AFLP标记对位于意大利maggiore湖地区的常绿杜鹃花品种分类的贡献。幼发拉底卡158，47-66。[doi: 10.1007/s10681-007-9425-3](https://doi.org/10.1007/s10681-007-9425-3)

Schoch，C. L .，Seifertb，K. A .，Huhndorfc，s .，Robertd，v .，Spougea，J. L .，Levesqueb，C. A .，et al. (2015)。核核糖体内部间隔区(ITS)区域作为真菌的通用DNA条形码标记。继续。纳特。阿卡德。Sci。U.S.A. 109，6241–6246。[doi:10.1073/PNA。18860 . 888888888816](https://doi.org/10.1073/pnas.1117018109)

夏普，P. M .，李，W. H. (1986)。大肠杆菌中调节基因的密码子使用并不反映对“稀有”密码子的选择。核酸第14，7737–7749号决议。[doi: 10.1093/nar/14.19.7737](https://doi.org/10.1093/nar/14.19.7737)

Shearman，J. R .，Sonthirod，c .，Naktang，c .，Pootakham，w .，Yoocha，t .，Sangsrakru，d .，et al. (2016)。甘蔗品种线粒体基因组的两条染色体:使用长PacBioreads的组装和重组分析。Sci。代表6:31533。[doi: 10.1038/srep31533](https://doi.org/10.1038/srep31533)

沈，J. S .，李，X. Q .，朱，X. T .，黄，X. L .，以及金，S. H. (2019).观赏药用和食用树木普勒杜鹃的完整叶绿体基因组。线粒体DNA B 4，3527–3528。[doi: 10.1080/23802359.20191676181](https://doi.org/10.1080/23802359.2019.1676181)

沈，J. S .，李，X. Q .，朱，X. T .，黄，X. L .，及金，S. H. (2020).杜鹃花质体全基因组及杜鹃花科植物的比较遗传分析。森林11:158。[doi: 10.3390/f11020158](https://doi.org/10.3390/f11020158)

Shirasawa，k .，Kobayashi，n .，Nakatsuka，a .，Ohta，h .，和Isobe，S. (2021)。两种杜鹃花的全基因组测序及分析。DNA Res. 28，1-7。[doi: 10.1093/dnares/dsab010](https://doi.org/10.1093/dnares/dsab010)

Sloan，D. B .、Alverson、A. J .、Chuckalovcak、J. P .、Wu、m .、McCauley、D. E .、Palmer、J. D .、et al. (2012)。开花植物线粒体中巨大的多染色体基因组快速进化，突变率极高。PLoS Biol。10:1001241。[doi:10.1371/journal . pbio . 1001241](https://doi.org/10.1371/journal.pbio.1001241)

Sloan，D. B .，Wu，z .，Sharbrough，J. (2018)。拟南芥参考线粒体基因组中持续性错误的纠正。植物细胞30，525–527。[doi:10.1105/tpc.18.00024](https://doi.org/10.1105/tpc.18.00024)

Smith，D. R. (2011)。将有限转移窗假说推广到细胞器间DNA迁移。Genome Biol。Evol。3, 743–748.[doi: 10.1093/gbe/](https://doi.org/10.1093/gbe/evr068)e[vr068](https://doi.org/10.1093/gbe/evr068)

Sugiura，M. (1995年)。叶绿体基因组。生物化学论文。30, 49–57.[doi:10.1007/978-94-011-2656-4\_10](https://doi.org/10.1007/978-94-011-2656-4_10)

孙，z .，万，D. G .，墨菲，R. W .，马，l .，张，X. S .，黄，D. W. (2009).昆虫线粒体基因组碱基组成和密码子使用的比较。Genes Genom。31, 65–71.[doi: 10.1007/BF03191139](https://doi.org/10.1007/BF03191139)

Thiel，t .，迈克莱克，w .，Varshney，R. K .，和Graner，A. (2003)。利用EST数据库开发和鉴定大麦(大麦)的基因衍生SSR标记。或者。Appl. Genet。106, 411–422.[doi:10.1007/s00122-002-1031-0](https://doi.org/10.1007/s00122-002-1031-0)

Thomas，W. A. B .、Michael，B. B .、马赛丽，s .和John，V. F. (2017)。边缘:非光合成杜鹃花科高度退化的质体。新植物醇。216:254.[doi: 10.1111/nph.14681](https://doi.org/10.1111/nph.14681)

Tillich，m .，Lehwark，p .，佩利泽，t .，Ulbricht-Jones，E. S .，Fischer，a .，Bock，r .，et al. (2017)。GeSeq——细胞器基因组的通用和精确注释。核酸Res. 45，W6-W11。[doi: 10.1093/nar/gkx391](https://doi.org/10.1093/nar/gkx391)

Timmis，J. N .，Ayliffffe，M. A .，Huang，C. Y .，Martin，W. (2004)。内共生基因转移:细胞器基因组形成真核染色体。纳特。热内特牧师。5，123–U116。[doi: 10.1038/nrg1271](https://doi.org/10.1038/nrg1271)

Walker，B. J .。Pilon:用于全面微生物变体检测和基因组组装改进的集成工具。PLoS One 9:e112963。[doi: 10.1371/journalpone.0112963](https://doi.org/10.1371/journal.pone.0112963)

王，L. Q .，张，h .，姜，m .，陈，h .，黄，l .，刘，C. (2019)。Icacinaceae中第一个完整的石菖蒲质体序列、比较基因组分析和对气候变化做出反应的石菖蒲种的可能分裂。PeerJ 7:e6663。[doi: 10.7717/peerj。6663](https://doi.org/10.7717/peerj.6663)

Wick，R. R .，Judd，L. M .，Gorrie，C. L .，Holt，K. E. (2017)。Unicycler:从短和长测序reads中解析细菌基因组集合。PLoS Comput。Biol。13:e1005595。[doi:10.1371/journal . pcbi . 1005595](https://doi.org/10.1371/journal.pcbi.1005595)

Wicke，s .，Schneeweiss，G. M .，DePamphilis，C. W .，Müller，K. F .，和Quandt，

D.(2011).陆生植物质体染色体的进化:基因含量、基因顺序、基因功能。植物模型。Biol。76, 273–297.[doi: 10.1007/s11103-011-](https://doi.org/10.1007/s11103-011-9762-4)9762-[4](https://doi.org/10.1007/s11103-011-9762-4)

f . Wright(1990)。基因中使用的“有效密码子数”。基因87，23–29。[doi:10.1016/0378-1119(90)90491-9](https://doi.org/10.1016/0378-1119(90)90491-9)

Wu，Z. Q .，Ge，S. (2012)。重新审视禾本科植物BEP分支的系统发育:来自叶绿体全基因组序列的证据。摩尔。系统发育。Evol。62, 573–578.[doi:10 . 1016/j . ym pev . 2011 . 10 . 019](https://doi.org/10.1016/j.ympev.2011.10.019)

Wu，Z. Q .，廖，X. Z .，张，X. N .，Tembrock，L. R .，Broz，A. (2022).植物线粒体的基因组结构变异--多染色体结构综述。J. Syst .Evol。60, 160–168.[doi: 10.1111/jse.12655](https://doi.org/10.1111/jse.12655)

Wu，z .，J. M .、Taylor，D. R .、Sloan，D. B. (2015)。被子植物夜花防风的大量线粒体基因组是通过整个染色体的获得或丧失而进化的。继续。纳特。阿卡德。Sci。U.S.A. 112，10185–10191。[doi:10.1073/PNA。18860 . 888888888816](https://doi.org/10.1073/pnas.1421397112)

徐，j，罗，h，聂，s，张，R. G，毛，J. F. (2021)。杜鹃花线粒体和质体基因组完整的杜鹃花，一个重要的亲本广泛种植的杜鹃花。线粒体DNA B 6，1197–1199。[doi: 10.1080/23802359.2021.1903352](https://doi.org/10.1080/23802359.2021.1903352)

张f、李w、高c、张d、高L. Z. (2019)。解读茶树叶绿体和线粒体基因组。阿萨米卡。Sci。数据6:209。[doi:10.1038/s 14598-018-35926-y](https://doi.org/10.1038/s41598-018-35926-y)

Zubaer，a .，Wai，a .，和Hausner，G. (2018)。树脂内缢管蚜线粒体基因组内含子丰富。Sci。代表8:17591。[doi:10.1038/s 14598-](https://doi.org/10.1038/s41598-018-35926-y)018-[35926-y](https://doi.org/10.1038/s41598-018-35926-y)