

东北师范大学生命科学学院  
2010 级野外实习自主研究性学习报告



# 东北师范大学

论文题目：基于叶绿体 DNA *trnL-F* 和 *rps16* 序列探讨  
长白山地区鳶尾科植物的亲缘关系

组 别：植物 第三组

组 长：李秋萍

组 员：李亚玲、张安麒、金贵花、窦润芝  
徐桂香、陈飞、汪强、初蕾

指导教师：孙明洲

## 基于叶绿体 DNA *trnL-F* 和 *rps16* 序列探讨 长白山地区鸢尾科植物的亲缘关系

**【摘要】**通过对鸢尾科的 12 种样品的 cpDNA 基因间隔区 *trnL-F* 和 *rps16* 的序列分析,研究部分鸢尾科植物的亲缘关系。运用 Clustal、PAUP 和 DNAMAN 软件对 12 种鸢尾科植物的 *trnL-F* 和 *rps16* 基因序列进行比对分析,构建分子系统树,可将 12 种物种分为 2 个组。通过研究叶绿体基因来确定鸢尾科植物的亲缘关系,为育种与改良提供依据。

**【关键字】**鸢尾科;叶绿体 DNA(cpDNA);亲缘关系; *trnL-F*; *rps16*

## Phylogenetic Relationships of Iridaceae Species in Changbai Mountains inferred from the Chloroplast DNA *trnL-F* and *rps16* sequence Data

**Abstract:** Phylogenetic relationships of 12 Iridaceae species inferred from *trnL-F* and *rps16* sequences data had been studied. *trnL-F* and *rps16*—genes belonging to 12 Iridaceae species were analyzed and compared to construct the phylogenetic trees using Clustal、PAUP and DNAMAN software. It showed that the twelve species were separated into two parts. Determining relationship of Iridaceae species through the study of chloroplast DNA, our study want to provide a basis for breeding and improvement.

**Key word:** Iridaceae; Chloroplast DNA(cpDNA); Phylogenetic relationship; *trnL-F*; *rps16*

## 前 言

鸢尾科植物属于被子植物门，目前公认的有 66 属约 2000 种，广泛分布于全世界的热带、亚热带及温带地区。我国产约 10 属，主要是鸢尾属植物，多分布于西南、西北及东北各地。

鸢尾科植物是多年生草本。常具根状茎、球茎或鳞茎。叶多基生，条形、剑形或为丝状。花两性，色泽鲜艳美丽，花或花序下有草质或膜质苞片，子房下位。果实为蒴果，常有附属物或小翅。本科植物以花大、鲜艳、花型奇特而著称，且栽培历史悠久，园艺品种及人工杂交种很多，花型及色泽变化也较大，深为各国园艺界所喜爱<sup>[1]</sup>。鸢尾科植物的应用价值主要体现在经济和药用上。东北是鸢尾的主要分布地，分布有大约 5 属 25 种<sup>[2]</sup>。其中长白山区有大约 10 种<sup>[3]</sup>。

### 1. 实验背景

#### 1.1 国内外研究进展

目前国际鸢尾科植物分类学研究正朝着系统植物学<sup>[4, 5]</sup>和进化植物学两个方向深入<sup>[6]</sup>。

国内对鸢尾的研究起步较晚，但在宋代已经有人把鸢尾作为药物使用<sup>[7]</sup>，《本草纲目》中就有对鸢尾属的马蔺的花色、花期的详细记载。我国最早的研究文章是 1936 年刘瑛发表的“中国之鸢尾”，记载了约三十种鸢尾<sup>[8]</sup>。对鸢尾科植物以分子生物学为手段的分类学研究也是在近些年开始的。目前的研究中，刘云 RAPD 技术对吉林产的 8 种鸢尾进行了聚类，得到了其亲缘关系图<sup>[9]</sup>。秦民坚等人通过对鸢尾、野鸢尾、蝴蝶花和德国鸢尾的叶绿体 *rbcL* 基因进行分析建立了

分子系统树<sup>[10]</sup>。牟少华等人基于叶绿体 DNA *trnL-F* 研究了 24 种鸢尾的亲缘关系，并构建了分子系统树<sup>[11]</sup>。

国外鸢尾科植物的研究起步较早。俄罗斯学者 Makarevitch 首次在 genebank 上批量提交了一些鸢尾属种 *trnL-F* 序列。Reeves 利用线粒体 DNA 四个间区序列对鸢尾属部分种进行了分子系统学研究<sup>[12]</sup>。Wilson 在 genebank 中提交了 20 个鸢尾属植物的种和亚种 ITS 序列，并用这 18 个 series Californicae 中种和亚种 ITS 序列，以 2 个 series Sibiricae 中的外类种群分析了系间和种内的系统关系<sup>[13]</sup>。南非的 Peter Goldblatt 是世界鸢尾科及其近缘种类的专家，著有若干专著，对鸢尾科的分类有很大贡献。

## 1.2 分类学研究

对于鸢尾科鸢尾属的分类一直是研究的热点。1817 年的 Roamer 和 Schultes 系统是鸢尾属植物最早建立的分类系统，这个系统将鸢尾属植物分为三大类群：（1）*Corollis barbatis*：在植株外花被裂片上有带色的多细胞的须毛；（2）*Corollis imberbibus*：外花被裂片上无须毛；（3）*Ominus cognitae*：除了以上两个群体外的全部种类。1823 年，Tausch 系统将鸢尾属下划分为 6 个组。1846 年 Spach 系统在鸢尾属下建立了 15 个亚属。1863 年，Alefeld 依据植株中较短的内花被、柱头分支及花被管的长度等特征来划分系统，将原来广义的属划分为许多独立的属。Bentham 和 Hooker 在 1883 以营养体为重要依据发表了一个新的系统。19 世纪末 Baker 所著的《鸢尾手册》收录了 160 多种鸢尾种类，并将鸢尾属植物划分为 10 个亚属。1913 年 Dykes 在

属下设立了 12 个组。1930 年 Dykes 发表了鸢尾属第一个自然分类系统,该系统全面、系统地对鸢尾属的属下等级进行了科学的分类。1953 年, Lawrence 系统在属下设立了 5 个亚属。1961 年, Rodionenko 发表的分系统, 在属下划分为 6 个亚属, 亚属下设组、亚组及系, 该系统的主要依据是植物体较为明显的特征, 如根部是否肥大呈纺锤形、外花被裂片的形态、中脉上有无附属物、附属物是须毛状还是鸡冠状等。1981-1989 年, Mathew 基于前人的研究也建立了鸢尾属下的分类系统, 是目前较为认可的<sup>[14]</sup>。

各个分类系统的划分依据不同, 但各有所长, 是研究鸢尾属植物的重要资料。但各个分类系统都是基于形态学特征来划分的。

### 1.3 *rps16* 和 *trnL-F* 基因特性

鸢尾是一种高等植物, 高等植物的核基因组进化速率十分快, 拥有许多具有丰富变异的 DNA 片段(如某些内含子区)<sup>[15]</sup>, 往往给进化、亲缘关系的研究带来困难。细胞质基因组它的特点是单亲遗传, 与核基因相比, 进化速率较慢。细胞质基因组包括线粒体 DNA (mtDNA) 和叶绿体 DNA (cpDNA)。线粒体基因的进化速度还不到叶绿体基因的 1/3, 而且普遍的重排现象使得线粒体基因组结构变化多样, 为 PCR 扩增和测序工作带来极大的不便, 因而 mtDNA 在植物系统发育研究中则应用较少<sup>[15]</sup>。叶绿体 DNA, 它分子量小、多拷贝、结构简单, 其 DNA 片段容易扩增和测序, 目前已被广泛用来研究植物的系统发育、分类等<sup>[16]</sup>。

叶绿体基因组编码区和非编码区序列进化速率相差较大<sup>[17]</sup>, 植物

叶绿体DNA (cpDNA) 在一些非编码区域存在相对较高的核苷酸置换率<sup>[18]</sup>, 从而使之在种间和种内都拥有较高的遗传变异<sup>[19]</sup>, 适用于在属一级水平或者属下水平的系统发育关系分析, 是迄今用于分子系统学和谱系地理学研究最多的一类基因组<sup>[20]</sup>。

为了研究长白山地区即中国高纬度带的鸢尾的亲缘关系, 本次选取了叶绿体的非编码区的 *rps16* 和 *trn L-F* 基因<sup>[17]</sup> 进行研究, 其中 *rps16* 是内含子基因, *trnL-F* 是 cpDNA 基因间区<sup>[21]</sup>。基于两个基因片段联合分析鸢尾科的亲缘关系可以弥补单基因建树的不足。

本次实验采用的是 DNA 序列分析, 它是在 DNA 一级结构水平上, 通过对比不同样品的核苷酸序列来研究植物的亲缘关系远近。

## 2 材料与方法

### 2.1 材料

本次研究所用的材料分别采自吉林、黑龙江、辽宁、内蒙古、新疆 (表 1)。吉林、黑龙江、辽宁三省东部山地总称长白山脉, 是中国的高纬度地区, 另选取纬度基本相同的新疆、内蒙古地区的一些样品作为比对。采集新鲜、完整的植物幼嫩叶片, 用硅胶迅速干燥, 直至完全变干后, 于超低温冰箱中保持待用。

| 编号<br>Code | 中名<br>Chinese name | 学名<br>Formal name                      | 材料来源<br>Sources |
|------------|--------------------|--|-----------------|
| B1         | 射干                 | <i>Belamcanda chinensis</i> (L.) DC.   | 吉林省白山市抚松县       |
| I1         | 溪荪                 | <i>I. Sanguinea</i> Donn ex Horn.      | 内蒙古根河市          |
| I2         | 紫苞鸢尾               | <i>I. ruthenica</i> Ker-Gawl.          | 河北张家口市蔚县        |
| I3         | 长尾鸢尾               | <i>I. rossii</i> Baker                 | 辽宁省丹东市五龙背       |
| I4         | 野鸢尾                | <i>I. dichotoma</i> Pall.              | 吉林省白城市平台镇       |
| I5         | 单花鸢尾               | <i>I. uniflora</i> Pall. ex Link       | 黑龙江黑河市          |
| I6         | 细叶单花鸢尾             | <i>I.uniflora</i> var. <i>caricina</i> | 内蒙古赤峰市          |
| I7         | 山鸢尾                | <i>I. setosa</i> Pall. ex Link         | 吉林省安图县二道白河镇     |
| I8         | 粗根鸢尾               | <i>I. tigridia</i> Bunge               | 内蒙古乌兰浩特市        |
| I9         | 膜苞鸢尾               | <i>I. scariosa</i> Willd. ex Link      | 新疆布尔津县喀纳斯       |
| I10        | 喜盐鸢尾               | <i>I. halophila</i> Pall.              | 新疆乌鲁木齐市         |
| I11        | 囊花鸢尾               | <i>I. ventricosa</i> Pall.             | 内蒙古乌兰浩特市        |

表 1 本研究中鸢尾科植物材料  
Tab.1 Information of Iridaceae species in this study



图 1 实验材料采样图

## 1.1 实验方法

### 1.2.1 DNA 提取

将样品用液氮充分研磨后，运用基因组 DNA 提取试剂盒（康为



世纪) 对所收集的样品进行总的 DNA 提取。

### 1.2.2 PCR 扩增及测序

PCR 扩增反应的总体积为 30  $\mu$  L, PCR 反应体系为 ddH<sub>2</sub>O 22.5  $\mu$  L, 10 $\times$ buffer, 0.3  $\mu$  L *Taq* DNA 聚合酶, BSA 0.6  $\mu$  L, 引物 1 和引物 2 分别 0.5  $\mu$  L, 以及 2  $\mu$  L 的模板 DNA。其中 *trnL-F* 引物来自 Taberlet<sup>[18]</sup>等发表的叶绿体通用引物, *rps16* 引物来自 Oxelman<sup>[22]</sup>等发表的叶绿体通用引物。反应程序为: 95 $^{\circ}$ C, 5min; 94 $^{\circ}$ C, 30s; 退火温度, 52 $^{\circ}$ C, 30s; 72 $^{\circ}$ C, 90s; 35 个循环; 72 $^{\circ}$ C, 8min; 4 $^{\circ}$ C 保存。实验所用 PCR 检测仪和毛细管荧光电泳仪型号分别为 ABI2720 (ABI, USA) 和 ABI3730 (ABI, USA)。PCR 扩增产物送至上海生工进行测序。

### 1.2.3 序列分析

测序所得的序列用 Clustal 进行序列比对, 并进行适当的人工校正。借助 PAUP Version4.0b10 软件, 采用 UPGMA (非加权组平均法) 构建了系统发育树, 此方法适合于分析遗传距离接近的材料。利用 DNAMAN (5.2.9 Demo version) 软件构建了同源关系树。

## 2 结果与分析

### 2.1 *rps16* 和 *trnL-F* 片段的序列变异

两种 PCR 扩增产物在琼脂糖凝胶上长度均为约 800bp。经序列测定 *rps16* 基因序列的长度为 817~852bp, 比较长度为 826bp, 有进化意义的位点共 208 个, 约占总序列比较长度的 25.18%, *trnL-F* 基因序列的长度为 839~868bp, 比较长度为 773bp, 有进化意义的位点 139 个, 约占总序列比较长度的 17.98%。

## 2.2 鸢尾科植物的亲缘关系分析

分别用 *rps16* 基因、*trnL-F* 基因以及编辑拼接了的两种基因构建了系统发育关系树，结果表明，鸢尾科植物都分成两个系，射干独成一系，鸢尾属下可分为两组，但是分别用 *rps16* 和 *trnL-F* 基因构建的系统树在具体分支上存在差异。

对于一个特定的类群，其进化历史是唯一的，即物种树是唯一的，当出现所构建的多个系统树之间出现冲突时，实质上是部分的系统树没能正确反映物种进化关系。这可能是由于随机误差、系统误差或者生物学因素<sup>[23]</sup>。为了减少以单个基因建树的误差，选择联合两种基因构建系统树，见图 1。

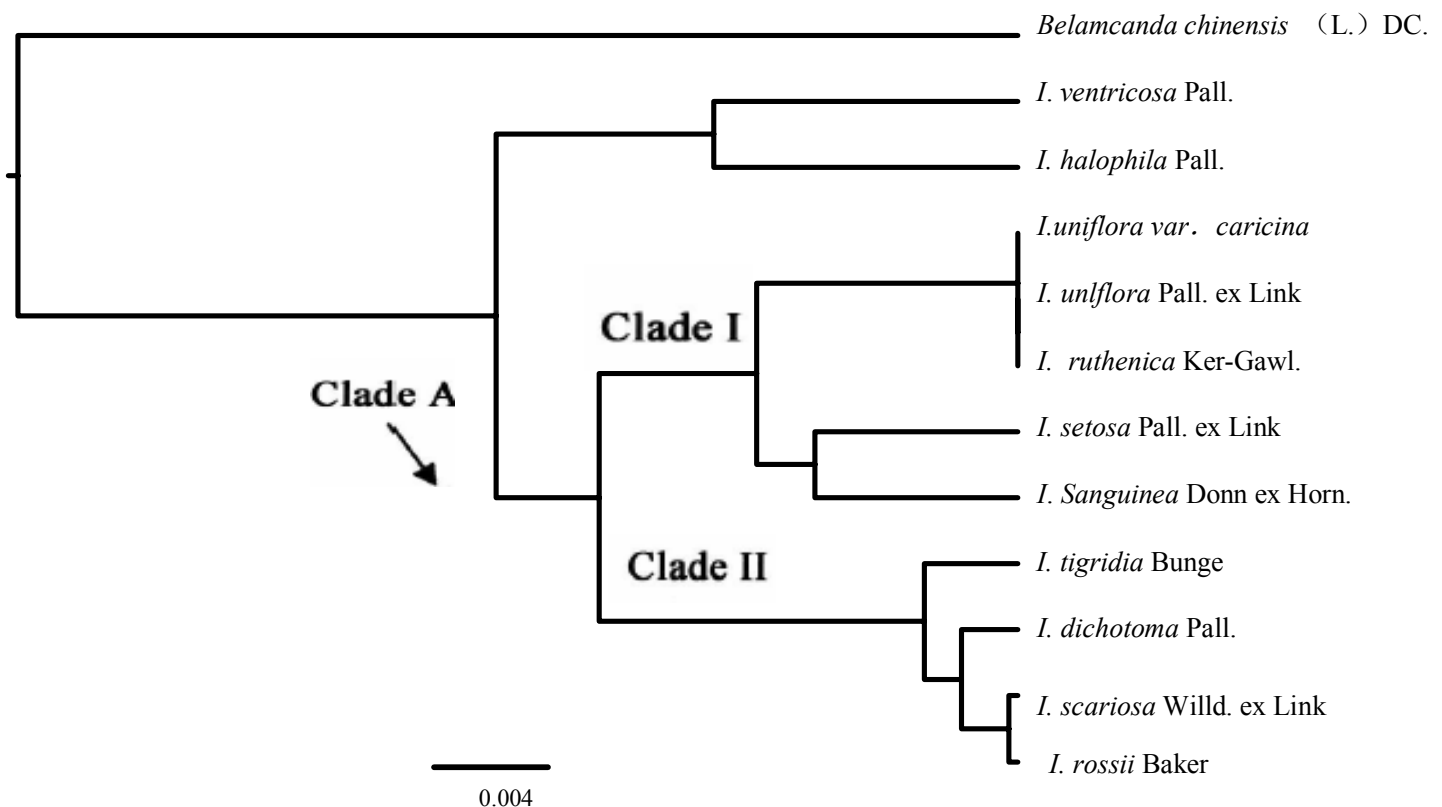


图 1 联合 *rps16* 和 *trnL-F* 基因构建的系统树

Fig.1 The phylogenetic tree reconstructed for Iridaceae based on the Chloroplast DNA *trnL-F* and *rps16* sequence Variation

总体上鸢尾科植物可以分为 3 个独立的分支，射干属的射干 B1 构成一个单系，鸢尾属被分成了 2 个大组。喜盐鸢尾 (I10) 和囊花鸢尾 (I11) 为一大组。长尾鸢尾 (I3) 与膜苞鸢尾 (I9) 关系很近，它们与野鸢尾 (I4) 聚在一起，和粗根鸢尾 (I8) 构成一组 Clade II；紫苞鸢尾 (I2)、单花鸢尾 (I5)、细叶单花鸢尾 (I6) 聚在一起，溪荪 (I1) 与山鸢尾 (I7) 聚在一起，它们共同聚成一组 Clade I，这 7 种鸢尾共同聚成一大组 Clade A。

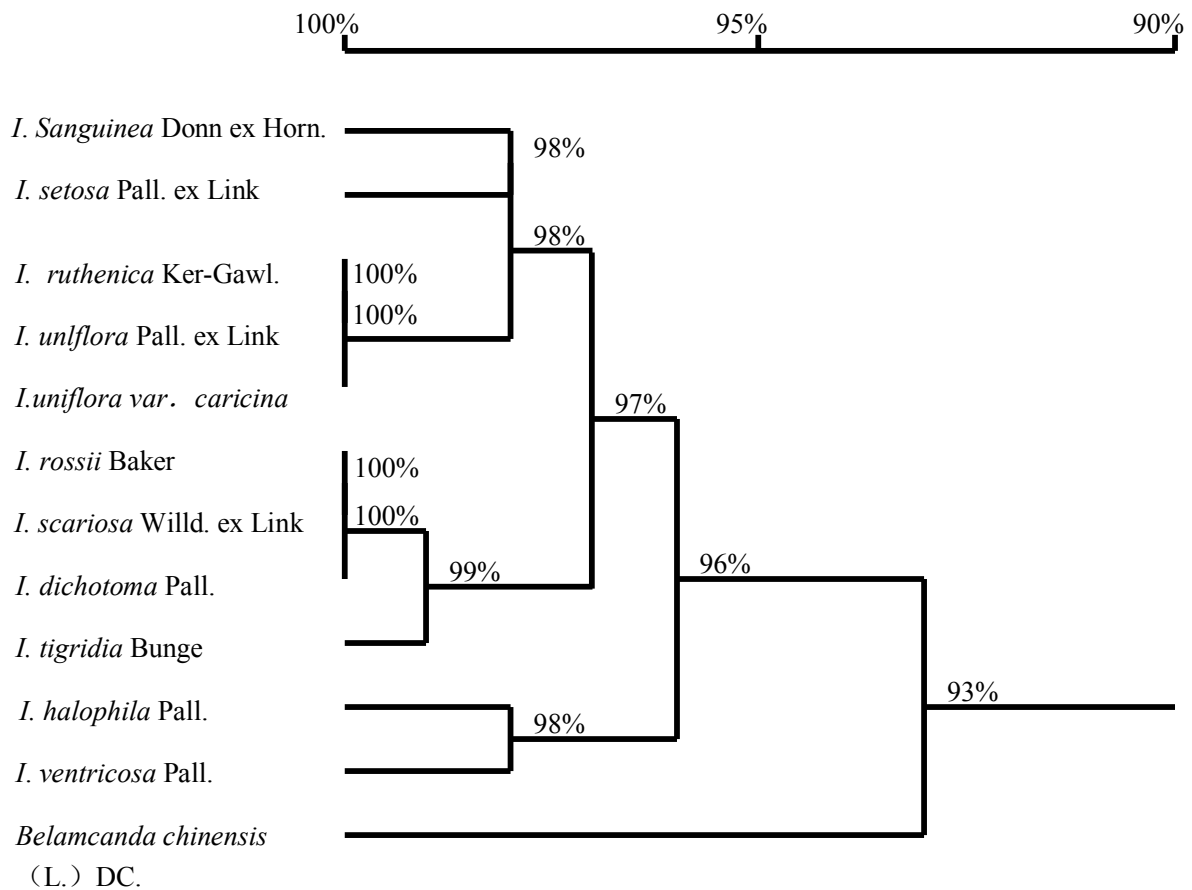


图 2 联合 *rps16* 和 *trnL-F* 基因构建的鸢尾科同源关系树

Fig.1 The homology-tree reconstructed for Iridaceae based on the Chloroplast DNA *trnL-F* and *rps16* sequence Variation

根据其同源关系树和系统发育树可以看出，所选择的长白山地区的鸢

尾科植物的同源性很高，遗传距离较小，其中射干（B1）与鸢尾属其它物种的遗传距离为 0.037，鸢尾属种内的遗传距离很小，表明其分化的年代并不是太久远。

### 3 讨论

#### 3.1 序列特性

筛选的两个鸢尾科植物的叶绿体基因具有较大的变异，其中 *rps16* 的变异位点占到总序列比较长度的 25.18%，*trnL-F* 基因序列的变异位点占到总序列比较长度的 17.98%。这可能是由于鸢尾属植物是一种古老的植物，起源于第三纪<sup>[24]</sup>。还可能的原因是所选的两个基因位于非编码区，所承受的自然选择的压力较小，变异被保留的机会很大<sup>[25]</sup>。这两种基因的变异程度很适于确定属内水平的关系，这也证明了所选基因的正确性。

本次实验选取了 *rps16* 基因和 *trnL-F* 基因，可以减少以一种基因建树的误差。

#### 3.2 长白山地区鸢尾科植物的亲缘关系

鸢尾科植物在中国分布广泛，本次研究所取材料主要是东北、西北等中国高纬度地区，以纬度基本相同的西北地区的几个种为对比，来探讨长白山地区的鸢尾科植物的亲缘关系。

研究结果表明射干属的射干与鸢尾属的物种的遗传距离较大，是符合形态学分类的。鸢尾属的长尾鸢尾、膜苞鸢尾、野鸢尾、粗根鸢尾聚为一支，溪荪、山鸢尾、紫苞鸢尾、单花鸢尾、细叶单花鸢尾、喜盐鸢尾、囊花鸢尾聚成一支。

我国的鸢尾属植物的分类是我国学者赵毓棠采用了 Rodioninko 的系统，对鸢尾属各个种的分类做了调整。它是以形态学特征为主要依据，包括根状茎类型、外花被裂片、中脉上有无附属物等。参照该系统，探讨存在的一些问题：

(1) UPGMA 聚类树自展支持率都在 93%。细叶鸢尾、单花细叶鸢尾和紫苞鸢尾聚成一支，它们的亲缘关系很近。在传统分类中，它们都属于无附属物亚属的紫苞鸢尾组。山鸢尾和溪荪聚为一支，它们在传统的分类上都是无附属物亚属的无附属物组。可见 UPGMA 聚类树明确的区分了鸢尾属的不同亚属、组间的关系。

(2) 无附属物亚属的分类较为复杂，它是鸢尾属中分布范围最广的一个亚属，种类繁多，属下分类等级也最多。在 UPGMA 聚类树中无附属物亚属的也较为分散。

#### 1. 喜盐鸢尾的位置关系

喜盐鸢尾属于琴瓣鸢尾亚属，该亚属在《中国植物志》中只有一个分类群，包括喜盐鸢尾和其变种喜盐蓝花鸢尾。喜盐鸢尾的分类在各种分类系统中存在明显分歧，Mathew 将它归为无附属物亚组中，而 Rodionenko 将其归为琴瓣鸢尾亚属。在 UPGMA 聚类树上看到，喜盐鸢尾与囊花鸢尾的亲缘关系较近，支持率为 98%。则中国将喜盐鸢尾划分到琴瓣鸢尾亚属中应当是错误的，应放入无附属物亚属。

(3) 早期研究将射干作为外类群进行讨论分析，随着形态、细胞、分子等研究的深入，多个研究证明它与野鸢尾亲缘关系较近，很多报道支持将两个种并在一起。在本次实验中，以 *trnL-F* 基因建立的

UPGMA 聚类树中它与野鸢尾的亲缘关系较近，但是以 *rps16* 基因以及结合两种基因建立的 UPGMA 聚类树显示射干应当作为一个独立的属。应当需要进一步的研究来得出结论。

由分子实验得出物种亲缘关系远近与传统的形态学划分有较大的出入。一方面这可能是由于叶绿体基因是单亲遗传并不能反映所有的系统发育问题，或者是所选择的基因序列并不够有代表性所致，还有可能是由于我们所选取的基因是叶绿体中进化速率较快的基因，但有效信息仍较少，可能无法反映出系统关系的全貌。另一方面可能是由于传统的形态学分类是基于植物的外观形态而进行的人为分类方法，本身存在一定的问题。现在的分子技术可以通过对物种的基因序列的比对得出物种在进化过程中的系统位置，可以用于修正传统分类学的错误。

国内外长期的研究实践证明，植物的结构特点和形态演化规律是研究植物系统与进化规律的重要性状和证据，应重视和分子生物学其它分支学科的知识和技术交叉、结合，共同为探讨植物系统与进化规律做出贡献<sup>[26]</sup>。

因此，为了更加准确的得到鸢尾科植物的亲缘关系，我们需要筛选出更合适的核基因或细胞质基因并增加研究的基因数量，与传统的形态性状结合起来共同分析，得到该科植物在系统发育中的正确关系，从而为鸢尾科植物的育种、改良、药用价值等研究等提供一定的依据。

## 参考文献

- [1] 赵毓堂. 中国植物志(第十六卷·第一分册)[M]. 北京:科学出版社, 1985:120-197.
- [2] 傅沛云, 孙启时, 陈佑安. 等. 东北草本植物志, 第12卷. 北京:科学出版社, 1998.
- [3] 涂英芳, 杨野, 衣俊鹏, 等. 长白山野生观赏植物. 北京:中国林业出版社, 1993.
- [4] John C.Manning, Peter Goldblatt. Endothecium in Iridaceae and Its Systematic Implications. *American Journal of Botany*, 1990, 77 (4) : 527-532.
- [5] Kron P; Stewen S C. Variability in the expression of a rhizome architecture model in a natural population of *Iris versicolor* (Iridaceae)[J]. *American Journal of Botany*, 1994, 81(09):1128-1138. DOI:10.2307/2445475.
- [6] Peter Goldblatt. Systematics phylogeny and Evolution of Diets (Iridaceae). *Annals of the Missouri Botanical Garden*, 1981, 68 (1) : 132-153.
- [7] 仲轶. 基于多基因组的鸢尾属部分种的系统分类研究[D]. 东北林业大学, 2010.
- [8] 刘瑛. 中国之鸢尾. *中国植物学杂志*, 1963, (2) : 929-956.

- [9] 刘云. 吉林省产八种鸢尾属(*Iris*L.)植物的 RAPD 分析及结构植物学分析[D]. 东北师范大学, 2001
- [10] 秦民坚, 黄芸, 杨光等. 射干及类似药用植物叶绿体 *rbcL* 基因序列分析[J]. 药学学报, 2003, 38(2):147-152.
- [11] 牟少华, 孙振元, 彭镇华等. 基于叶绿体 DNA *trnL-F* 序列研究部分鸢尾属的亲缘关系[J]. 华北农学报, 2010, 25(5):112-116.
- [12] G. Reeves, P. Goldblatt, P. J. Rull. Molecular Systematics of Iridaceae: A Combined analysis of four plastid DNA sequence matrices [J]. *Annali di Botanica*, 2012, 58(1).
- [13] Wilson, Carol Anne. Evolutionary Studies in Iris Series Californicae. PhD. University of California, Berkeley, 1996, 116.
- [14] 王玲. 鸢尾属部分种发育生物学与系统演化的研究[D]. 东北林业大学, 2005
- [15] 方伟, 杨俊波, 杨世雄; 李德铤. 基于叶绿体四个 DNA 片段联合分析探讨山茶属长柄山茶组、金花茶组和超长柄茶组的系统位置与亲缘关系[J]. 云南植物研究, 2010, 32(01): 1-13
- [16] 汪小全, 洪德元. 植物分子系统学近 5 年的研究进展概况, 植物分类学报, 1997, 35(5): 465-480.
- [17] 高巾帼, 田兴军, 陈彬等. DNA 分子标记在植物系统学中的应用. 中国生物多样性保护与研究进展. 216-228.
- [18] Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant*



Molecular Biology 17: 1105–1109.

[19] Newton A C, Allnot T R, Gillies A C M, Lowe A J, Ennos R A. 1999. Molecular phylogeography, intraspecific variation and conservation of tree species. *Trends in Ecology and Evolution* 14: 140–145.

[20] 张茜, 杨瑞, 王钦等. 基于叶绿体 DNA *trnT-trnF* 序列研究祁连圆柏的谱系地理学[J]. *植物分类学报*, 2005, 43(6): 503–512.

[21] 田欣, 李德铎. DNA 序列在植物系统学研究中的应用[J]. *云南植物研究*, 2002, 24(2): 170–184.

[22] Oxelman B, Liden M, Berglund D. 1997. Chloroplast *rps16*

[23] 邹新慧, 葛颂. 基因树冲突与系统发育基因组学研究[J]. *植物分类学报*, 2008, 46(6): 795–807. DOI: 10.3724/SP.J.1002.2008.08081.

[24] 谢航. 中国鸢尾属有关分类群的讨论及属下分类系统的修订[D]. 东北师范大学, 1996.

[25] Demesore B, Sodzi N, Petit R P. A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants [J]. *Mol Eol*, 1995, 4: 129–131.

[26] 胡正海. 植物比较解剖学在中国 50 年的进展和展望[J]. *西北植物学报*, 2003, 23(2): 344–355.

附录:

CLUSTAL X (1.81) multiple sequence alignment

```

I3      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGGCA
I9      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGGCA
I4      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGGCA
I1      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGACA
I7      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGACA
I2      GAGTACTTGT----ATCTATGTAAATTAAGAGACTGCCAAAGCATTAAACAAATTTGACA
I5      GAGTACTTGT----ATCTATGTAAATTAAGAGACTGCCAAAGCATTAAACAAATTTGACA
I6      GAGTACTTGT----ATCTATGTAAATTAAGAGACTGCCAAAGCATTAAACAAATTTGACA
I10     GAGTACTTGTGGTATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGACA
I11     GAGTACTTGTGGTATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGACA
I8      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGGCA
B1      GAGTACTTGT----ATCTATGTAAATTAAGAGACTACCAAAGCATTAAACAAATTTGGCA
*****

```

```

I3      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I9      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I4      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I1      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I7      CAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I2      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I5      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I6      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
I10     TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTAAAAAGGAAGACTTTCTTTGTAGTA
I11     TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTAAAAGGAAGACTTTCTTTATAGTA
I8      CAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
B1      TAGTCCCTGTAATTTCTTAGTTTTAGATCTTCATTCAAAAAGAAGACTTTCTTT---GTA
***

```

```

I3      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
I9      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
I4      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
I1      AACGTAAGGATAATTATACGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
I7      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
I2      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATGC
I5      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATGC
I6      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATGC
I10     AACATAAGGATAATGAGATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
I11     AACGTAAGGATAATGATATGGACTGTAAATGATTCAATAATGGGAATCACTTACCCATAC
I8      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC
B1      AACGTAAGGATAATGATATGGACTGTGAATGATTCAATAATGGGAATTACTTGCCCATAC

```



I3 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I9 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I4 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I1 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I7 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I2 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I5 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I6 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I10 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I11 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 I8 ATAGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT  
 B1 ATGACTTTTTATTGGGGATAGAGGGACTTGAACCCTCACGATTTCTAAAGTCGACGGAT

\*\*\*\*\*

I3 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAAAATGGGACTCTCTCT  
 I9 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAAAATGGGACTCTCTCT  
 I4 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAAAAGGGACTCTCTCT  
 I1 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 I7 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 I2 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 I5 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 I6 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 I10 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 I11 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAAAGGGACTCTCTCT  
 I8 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT  
 B1 TTTCTTTTACTATAAAATTCATTGTTGTCGGTATTGACATGTAGAATGGGACTCTCTCT

\*\*\*\*\* \*\* \*\*\*\*\*

I3 TTATTCTTTCTGATTAATTCATTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I9 TTATTCTTTCTGATTAATTCATTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I4 TTATTCTTTCTGATTAATTCATTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I1 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I7 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I2 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I5 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I6 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I10 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I11 TTATTCTTTATGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 I8 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT  
 B1 TTATTCTTTCTGATTAATTCGTTTTCAAAGATCTATCAAACCTCTGGAATGAATGATT

\*\*\*\*\* \*\*\*\*\* \*\*\*\*\*

I3 TGATCACTGATCGATTCCCTCTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I9 TGATCACTGATCGATTCCCTCTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT

I4 TGATCACTGATCGATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I1 TGATCACTGATCGATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I7 TGATCACCGATCAATTCTCCTTCAACTTCGATTGGAATAGGTCCACAATAGCTCTGAAT  
 I2 TGATCACTGATCAATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I5 TGATCACTGATCAATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I6 TGATCACTGATCAATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I10 TGATCACTGATCGATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 I11 TGATCACCGATCGATTCTCCTTCAACTTCGATTGGAATAGATCCAC-----  
 I8 TGATCACTGATCGATTCTCCTTCAACTTCGATTGGAATAGATCCACAATAGCTCTGAAT  
 B1 TGATCACTGATCGATTCTCCTTCAACTCCGATTGGAATAGATCCACAATAGCTCTGAAT  
 \*\*\*\*\* \*\*

I3 TTTGCATATATTTGCATATATAATGATAATGGATTCAGGTCGTGATTAATCATTGATTA  
 I9 TTTGCATATATTTGCATATATAATGATAATGGATTCAGGTCGTGATTAATCATTGATTA  
 I4 TTTGCATATATTTGCATATATAATGATAATGGATTCAGGTCGTGATTAATCATTGATTA  
 I1 TTTGCATATAT-----AATGGATTCAGGTCGTGATTAATCGTTTGACTA  
 I7 TTTGCATATAT-----AATGGATTCAGGTCGTGATTAATCGTTTGATTA  
 I2 TTTGCATATAT-----AATGGATTCAGGTCGTGATTAATCGTTTGATTA  
 I5 TTTGCATATAT-----AATGGATTCAGGTCGTGATTAATCGTTTGATTA  
 I6 TTTGCATATAT-----AATGGATTCAGGTCGTGATTAATCGTTTGATTA  
 I10 TTTGCATATAT-----AATGGATTCAGGTCGTGATTAATCGTTTGATTA  
 I11 -----TAT-----AATGGATTCAGGTCGTGATTAATCGTTTGATTA  
 I8 TTTGCATATAT-----AATGATAATGGATTCAGGTCGTGATTAATCATTGATTA  
 B1 TTTGCATATAT-----AATGATAATGGATTCAGGCCGTGATTAATCGTTTGATTA  
 \*\*\* \*\*\*\*\* \*\*

I3 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTTGATTCTAGTTTTAAACCCTGT  
 I9 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTTGATTCTAGTTTTAAACCCTGT  
 I4 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTTGATTCTAGTTTTAAACCCTGT  
 I1 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTTATTCTAGTTTTAAACCCCTGT  
 I7 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTT-ATTCTAGTTTTAAACCCCTGT  
 I2 GAACAGCTTCCGTTGAGTCTCTGCACCTATCCTTTTTATTCTAGTTTTAAACCCCTCGT  
 I5 GAACAGCTTCCGTTGAGTCTCTGCACCTATCCTTTTTATTCTAGTTTTAAACCCCTCGT  
 I6 GAACAGCTTCCGTTGAGTCTCTGCACCTATCCTTTTTATTCTAGTTTTAAACCCCTCGT  
 I10 GAACAGCTTCCATTGAGTCTCTGCACCTATCCCCTTTTTATTCTAGTTTTAAACCCCTGT  
 I11 GAATAGCTTCCGTTGAGTCTCTGCACCTATCCCCTTTTTATTCTAGTTTTAAACCCCTGT  
 I8 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTTGATTCTAGTTTTAAACCCTGT  
 B1 GAACAGCTTCCATTGAGTCTCTGCACCTATCCTTTTTGATTCTAGTTTTAAACCCTGT  
 \*\*\* \*\*\*\*\* \* \*\*\*\*\* \*\*

I3 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCTAGAAAACG  
 I9 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCTAGAAAACG  
 I4 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCTAGAAAACG  
 I1 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCTAGAAAACG

I7 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 I2 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 I5 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 I6 TCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 I10 TTGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 I11 CCGT-----TTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 I8 TCGTGTTCGTTTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG  
 B1 TCGTGTTCGTTTTCTCAAATAAAAGATTTGGCTCAGGATTGCCCCCCCCCCTAGAAAACG

\*\* \*\*\*\*\*

I3 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I9 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I4 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I1 TACAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I7 TACAGGAGGTTTTCTCCGCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I2 TATAGGAGGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I5 TATAGGAGGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I6 TATAGGAGGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I10 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I11 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 I8 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT  
 B1 TATAGGAGGTTTTCTCCTCATAACGGCTCGAGAAAAAATGATTCTAATTTCTGTGTATAAT

\*\* \*\*\*\*\*

I3 AGCAAATGGATACATAAGAGCATGAATTAGACTATGATTTTAGTTATTTTTTTGTTCTTC  
 I9 AGCAAATGGATACATAAGAGCATGAATTAGACTATGATTTTAGTTATTTTTTTGTTCTTC  
 I4 AGCAAATGGATACATAAGAGCATAAATTAGACTATGATTTTAGTTATTTTTTTGTTCTTC  
 I1 AGCAAATGGATACATAAGAGCATGGATTAGACTATGATTTTAGTTATTTTTTTGTTCTTC  
 I7 AGCAAATGGATACATAAGAGCATGAATTAGACTATGATTTTAGTTATTTTTTTGTTCTTC  
 I2 AGCAAATGGATATATAAGAGCATGAATTAGACTACGATTTTAGTTATTTTTTTGTTCTTC  
 I5 AGCAAATGGATATATAAGAGCATGAATTAGACTACGATTTTAGTTATTTTTTTGTTCTTC  
 I6 AGCAAATGGATATATAAGAGCATGAATTAGACTACGATTTTAGTTATTTTTTTGTTCTTC  
 I10 AGCAAATGTATACATAAGAACATGAATTAGACTATGATTTTAGTTCTTTTTTTT-TTCTTC  
 I11 AGCAAATGTATACATAAGAACATGAATTAGACTATGATTTTAGTTCTTTTTTTT-TTCTTC  
 I8 AGCAAATGGATACATAAGAGCATGAATTAGACTATGATTTTAGTTCTTTTTTTGTTCTTC  
 B1 AGCAAATGGATACATAAGAGCATGAATTAGACTATGATTTTAGTTATTTTTTTGTTCTTC

\*\*\*\*\* \*\* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\* \*\*\*\*\*

I3 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATTC  
 I9 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATTC  
 I4 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATTC  
 I1 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATGC  
 I7 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATGC  
 I2 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATGT

I5 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATGT  
 I6 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATGT  
 I10 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATTC  
 I11 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATTC  
 I8 ACTACACTTACAGAAAAAAGAAATATCATTGTACTCATAACTCAAGTTGGATAATTC  
 B1 ACTACACTTACAGAAAAAAGAAATATCATTGTACCCAAAACTCAGTTGGAAAAATTC  
 \*\*\*\*\*

I3 GTAAAGAATTCAAAGTGAAATCCTCAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I9 GTAAAGAATTCAAAGTGAAATCCTCAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I4 GTAAAGAATTCAAAGTGAAATCCTCAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I1 GTAAAGAATTCGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I7 GTAAAGAATTCGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I2 GTAAAGAATTCGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTC-----  
 I5 GTAAAGAATTCGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTC-----  
 I6 GTAAAGAATTCGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTC-----  
 I10 GTAAAGAATTTGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I11 GTAAAAATTTGAAGTGAAATCCTTAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 I8 GTAAAGAATTCAAAGTGAAATCCTCAGAAATTTATTGAGTCGTCTCTAACCTCTTTTGTT  
 B1 GAAAAAATTCAAAGGGAAACCCTTAAAAATTTATGGAGCCGCCCTAACCCCTTTTGTT  
 \* \*\* \* \*\* \*

I3 TGTATCATCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I9 TGTATCATCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I4 TGTATCATCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I1 TGTATCGTCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I7 TGTATCGTCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I2 -----GAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGATACAATTGATTA  
 I5 -----GAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGATACAATTGATTA  
 I6 -----GAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGATACAATTGATTA  
 I10 TGTATCGTCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I11 TGTATCGTCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 I8 TGTATCATCTCGAATCTATTTTTTTCCTCATTCTAATAAAAATTATTGAGACAATTGATGA  
 B1 TGTATCACCCAAATCTTTTTTTTCCCATCTAATAAAAATTATTGAAACAATTGAGGA  
 \*\*\*\*\*

I3 AAATTGCGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAG  
 I9 AAATTGCGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAG  
 I4 AAATTGTGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAG  
 I1 AAATTGTGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAA  
 I7 AAATTGTGTTTCCTTGTTCGGAATCCCGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAA  
 I2 AAATTGTGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAG  
 I5 AAATTGTGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAG  
 I6 AAATTGTGTTTCCTTGTTCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTTAG

I10 AAATTGCGTTTCCTTGTTCTGAAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTAG  
 I11 AAATTGCATTTCCCTGTTATGAAATCCCGTCTCTTTGCTTTGAGAAATCCTTGGGTTAG  
 I8 AAATTGTGTTTCCTTGTTCCGGAATCCTGTCTCTTTGCTTTGTGAAATCCTTGGGTTAG  
 B1 AAATTGCGTTTCCTTGTTCCGGAACCCGGTCCCTTTGCTTTGGGAAACCCTGGGTTAA  
 \*\*\*\*\* \*\* \*\* \*\* \* \*\* \*\* \*\* \*\*\*\*\* \*\* \*\* \*\*\*\*\*

I3 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTT-----  
 I9 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTT-----  
 I4 ACATTACTTCGGCGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTT-----  
 I1 ACATTACTTCGGCGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I7 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I2 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I5 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I6 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I10 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I11 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 I8 ACATTACTTCGGTGATTCTTACTTCTTTCAAAATGGCAGCAACATACCTTTTGTTTTTTG  
 B1 ACATTACTTCGGGGATTCTTACTTCTTTCAAAATGGCACCAACATACCTTTTGTTTTTTG  
 \*\*\*\*\* \*\*\*\*\* \*\*\*\*\*

I3 --ATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I9 --ATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I4 --ATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I1 TTATTTATTTCTATGGAAAAGAAATACGAAGGATTGATTCCTTTGTAATA-----CACT  
 I7 TTATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I2 TTATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I5 TTATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I6 TTATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 I10 TTATTTATTTCTTTCCACAGAAATATGAAGGATTGATTCCTTTGTAATC-----CACT  
 I11 TTATTTATTTCTGTGAAAAGAAATATGAAGGATTGATTCCTTTGCAATTGTAATACACT  
 I8 TTATTTATTTCTATGGAAAAGAAATATGAAGGATTGATTCCTTTGTAATA-----CACT  
 B1 TTATTTATTTCTATGAAAAAATAGGAGGATTGATTCCTTTGTAATA-----CACT  
 \*\*\*\*\* \* \* \* \*\*\*\*\* \*\* \*\*\*\*\* \*\*\*\*\*

I3 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTGACTTTATTTTGAATTCAAACTTG  
 I9 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTGACTTTATTTTGAATTCAAACTTG  
 I4 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTGACTTTATTTTGAATTCAAACTTG  
 I1 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG  
 I7 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG  
 I2 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG  
 I5 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG  
 I6 TTACATCGAAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG  
 I10 TTACATCAAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG  
 I11 TTACATCTAAAAGTTTTCCAAATCCGACAAATTTACTTTATTTTGAATTCAAACTTG



```

I8          TTACATCGAAAAAGTTTCCAAATTCGACAAATTTGACTTTATTTTGAATTCAAAGTTG
B1          TTACATCGAAAAAGTTTCCAAATTCGACAAATTTGACTTTATTTTGAATTCAAAGTTG
          *****
I3          TTCGAATTTGAACCTTTCAATTTATATATTGATATTGAGGATATACTTACAAAGTTAGTC
I9          TTCGAATTTGAACCTTTCAATTTATATATTGATATTGAGGATATACTTACAAAGTTAGTC
I4          TTCGAATTTGAACCTTTCAATTTATATATTGATATTGAGGATATACTTACAAAGTTAGTC
I1          TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I7          TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I2          TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I5          TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I6          TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I10         TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I11         TTCGAATTTGAACCTTTCAATTTATATATTGA-----GGATATACTTACAAAGTTGGTC
I8          TTCGAATTTGAACCTTTCAATTTCTATATTGATATTGAGGATATACTTACAAAGTTAGTC
B1          GTCGAATTTGA-CCTTCTTTTATATATTGA-ATTGAGGATATACTTACAAAGTAGTAT
          ** *** * ***** ** ***** *****
I3          TCACTTATTCATTGTTACTAACCCTAGATTTTGCCCCCGATAAATGAATAAA-----GA
I9          TCACTTATTCATTGTTACTAACCCTAGATTTTGCCCCCGATAAATGAATAAA-----GA
I4          TCACTTATTCATTGTTACTAACCCTAGATTTTGCCCCCGATAAATGAATAAA-----GA
I1         TCACTTATTCATTGTTACTAACCCTAGATTTTTCCCCCGATAAATGAATAAAATAAAAGA
I7          TCACTTATTCATTGGTACTAACCCTAGATTTT-CCCCCGATAAATGAATAAAATAAAATA
I2          TCACTTATTCATTGTTACTAACCCTAGATTTTTCCCCCGATAAATGAATAAAATAAAAGA
I5          TCACTTATTCATTGTTACTAACCCTAGATTTTTCCCCCGATAAATGAATAAAATAAAAGA
I6          TCACTTATTCATTGTTACTAACCCTAGATTTTTCCCCCGATAAATGAATAAAATAAAAGA
I10         TCACTTATTCATTGTTACTAACCCTAGATTTTC-CCCCCGATAAATGAATAAAATAAAAGA
I11         TCACTTATTCATTGTTACTAACCCTAGATTTTC-CCCCCGATAAATGAATAAAATAAAAGA
I8          TCACTTATTCATTGTTACTAACCCTAGATTTTGCCCCCGATAAATGAATAAA-----GA
B1          CCACCTAT-CATTG-TACTA-CCCTAGATTTTGCCCCCGATAA-TGAATAA-----GA
          ***** ***** ***** ***** ***** ***** *
I3          TTTTTACTCGAGCTCCATCATGTACTATTTTATTTACCAACCAATACAAATTAGGTT
I9          TTTTTACTCGAGCTCCATCATGTACTATTTTATTTACCAACCAATACAAATTAGGTT
I4          TTTTTACTCGAGCTCCATCATGTACTATTTTATTTACCAACCAATACAAATTAGGTT
I1         TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCAATACAAATTGGGTT
I7          TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCAATACAAATCGGGTT
I2          TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCAATACAAATTAGGTT
I5          TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCAATACAAATTAGGTT
I6          TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCAATACAAATTAGGTT
I10         TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCAATACAAATTAGGTT
I11         TTTTTACTCGAGCTCCATCATGTACTATTTCTATTTACCAACCTAATACAAATTAGGTT
I8          TTTTTACTCGAGCTCCATCATGTACTATTTTATTTACCAACCAATACAAATTAGGTT
B1          TTTTTACTCGAGCT-ATCATGT-CTATTTT--ATTACCAACC--ATA-AAATTAGGTT
    
```

\*\*\*\*\*  
 \*\*\*\*\*  
 \*\*\*\*\*  
 \*\*\*\*\*  
 \*\*\*\*\*

I3 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATG  
 I9 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATG  
 I4 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATG  
 I1 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATA  
 I7 CTTAGCAGGAACAGAACAACAACTATGTCGAGTAAAGAGCATCTTCATTCCTATAGAAAATA  
 I2 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATA  
 I5 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATA  
 I6 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATA  
 I10 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATA  
 I11 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGTATCTTCATTCCTATAGAAAATA  
 I8 CTTAGCAGGAACAGAACAACAACTATGTCGAGTCAAGAGCATCTTCATTCCTATAGAAAATG  
 B1 CT-AGCAGGACGACAAC--ATCTCGTCAAGC---ACTCATTCTAGAATGTGATGAATCCA

\*\* \*\*\*\*\* \* \*\*\* \* \*\*\* \*\* \*\* \* \*\*

I3 GTGGACGTAAAAA  
 I9 GTGGACGTAAAAA  
 I4 GTGGACGTAAAAA  
 I1 GTGGATGGAAAAA  
 I7 GTGGATGGAAAAA  
 I2 GTGGATGGAAAAA  
 I5 GTGGATGGAAAAA  
 I6 GTGGATGGAAAAA  
 I10 GTGGATGGAAAAA  
 I11 GTGGATGGAAAAA  
 I8 GTGGACGTAAAAA  
 B1 GCGATCAGTCTCA

\* \* \*

