

# 苏姜猪线粒体基因组全长测序 及线粒体基因多样性分析

王 洁<sup>1</sup>, 郭浩君<sup>3</sup>, 张晨岭<sup>2,3</sup>

(1.江苏农牧科技职业学院,江苏 泰州 225300)

(2.南京农业大学动物科技学院,江苏 南京 210095)

(3.江苏第二师范学院生命科学与化学化工学院,江苏 南京 210013)

[摘要] 从分子水平上探究苏姜猪的遗传多样性,采用二代测序方法测定了苏姜猪线粒体基因组全序列,整合 NCBI 数据库中已公布的家猪类,以及本实验室测定的苏姜猪线粒体基因组全序列,使用蛋白编码基因和 rRNA 基因联合的数据集,基于贝叶斯法(BI法)和最大似然法(ML法)构建了家猪类的系统发生树,测序结果表明,该样本线粒体基因组全长 15 436 bps,包含典型的家猪类线粒体基因组 37 条基因,结构紧凑,其线粒体基因组碱基组成(A+T)约占 40%,(C+G)约占 60%,tRNA 总共有 8 处错配,结果表明苏姜猪具有丰富的遗传多样性,江海型线粒体核苷酸多态性高于全国 6 个类型猪的平均水平,体现了其独特的遗传多样性.在基于线粒体基因组完整编码区的系统发育分析中,苏姜猪在中国 6 个类型家猪中单独成为一支,支持其作为一个新猪种而存在.

[关键词] 苏姜猪,线粒体基因组,遗传多样性,系统发育

[中图分类号] S828.8 [文献标志码] A [文章编号] 1001-4616(2019)04-0097-06

## Whole Genome Sequencing and Genetic Diversity of Mitochondrial Genome of Sujiang Pig

Wang Jie<sup>1</sup>, Guo Haojun<sup>3</sup>, Zhang Chenling<sup>2,3</sup>

(1.Jiangsu Vocational College of Agriculture and Animal Husbandry, Taizhou 225300, China)

(2.College of Animal Science and Technology, Nanjing Agricultural University, Nanjing 210095, China)

(3.College of Life Science, Chemistry and Chemical Engineering, Jiangsu Second Normal University, Nanjing 210013, China)

**Abstract:** To discover the genetic diversity of Sujiang pig at molecular level, the complete mitochondrial genome of Sujiang pig was sequenced by next generation sequencing. The Sujiang pig had high genetic diversity in 15436 bps 37 genes, which A+T estimated percentage of 40%, C+G estimated percentage of 60% of its mitochondrial genome and tRNA had 8 mismatches. The Jianghai type mitochondrial nucleotide polymorphism was higher than the average level of the all six pig types, which revealed its unique genetic diversity. Phylogenetic analysis based on the complete coding region of the mitochondrial genome using NJ method, Sujiang pig was a separate pig among the six types of domestic pigs in China, and support its unique pig pedigree.

**Key words:** Sujiang pig, mitochondrial genome, genetic diversity, phylogeny

猪在分类学上,属偶蹄目(*Artiodactyla*),猪科(*Suidae*),野猪属(*Sus*),野猪种(*Scrofa*).我国拥有丰富的地方猪资源,也是世界上家猪驯化较早的中心之一.但近 40 年,地方猪品种受到国外引进猪种的强烈冲击,在数量和规模上均显著降低,遗传资源受到威胁,同时受到迁徙等活动的影响,种质特性与品种名称也发生了明显变化.随着市场经济的发展,商品猪种养殖逐渐取代地方传统猪种的养殖,地方猪品种资源没有得到充分的研究和保护利用,部分品种已濒临灭绝甚至已经消失.而本文研究的苏姜猪是以姜曲海猪和杜洛克猪为育种素材,主要分布于江苏省等地区,历时 17 年 6 个世代选育而成,是优质瘦肉型猪的新品种.了解苏姜猪的基因多样性,对利用和开发这一新品种有着非常重要的意义<sup>[1-2]</sup>.

收稿日期:2019-03-19.

基金项目:国家发改委资助项目(11710015021).

通讯联系人:王洁,讲师,研究方向:动物遗传多样性. E-mail:dearwj1980@163.com

随着 DNA 分子技术的快速发展,分子标记手段已经能够很好地解决物种的分类以及地位,本文涉及的第二代测序技术其核心思想是边合成边测序(*Sequencing by Synthesis*),即通过捕捉新合成的末端标记来确定 DNA 的序列,其特点是高通量,同时在全基因组覆盖率高时,准确率也极高,结果可靠,是目前最为流行的基因组研究手段. 由于线粒体基因组具有结构简单、进化速率快及母系遗传等特征,能从基因排序和基因序列两个层次提供进化信息,因此线粒体基因组可以作为研究 DNA 复制和转录的良好模型,近年来也被广泛地运用到生物学研究领域. 线粒体基因组的研究方法在近几十年中不断发展,尤其最近高通量测序技术的出现,为线粒体全基因组测序提供了新的策略. 线粒体基因组一定程度上反映了动物的进化历程,是遗传和进化上广泛使用的分子标记,也是研究苏姜猪基因多样性以及系统发育的理想材料,可以很好地解决家猪内部的一些系统学问题.

本文采用第二代测序技术对苏姜猪的线粒体全基因组进行测序,得到了完整的线粒体基因组信息,再通过生物信息学的分析方法对苏姜猪线粒体基因组的多样性和系统发育地位进行分析,发现其独特的遗传多样性和特殊的进化地位.

1 材料与方法

抽取苏姜猪的前腔静脉血 3 mL 加入 0.3 mg 的肝素钠抗凝,混匀后迅速放入干冰中保存,血液样本送至诺禾致源公司进行总 DNA 提取和后续实验. 总 DNA 在 Illumina Hiseq 2000 平台进行测序.

1.1 二代测序流程

(1)文库制备,将 DNA 用雾化或超声波随机片段成几百碱基或更短的小片段. 用聚合酶和外切核酸酶把 DNA 片段切成平末端,紧接着磷酸化并增加一个核苷酸黏性末端. 然后将 Illumina 测序接头与片段连接.

(2)簇的创建,将模板分子加入芯片用于产生克隆簇和测序循环. 芯片有 8 个纵向泳道的硅基片. 每个泳道内芯片表面有无数的被固定的单链接头. 上述步骤得到的带接头的 DNA<sup>o</sup>片段变性成单链后与测序通道上的接头引物结合形成桥状结构,以供后续的预扩增使用. 通过不断循环获得上百万条成簇分布的双链待测片段.

(3)测序,分三步:DNA 聚合酶结合荧光可逆终止子,荧光标记簇成像,在下一个循环开始前将结合的核苷酸剪切并分解.

(4)得到的 raw data 在过滤后形成 clean data 进行分析.

1.2 线粒体基因组组装与基因注释

使用 Geneious 11.2<sup>[3]</sup>,以 GenBank 上已知家猪的线粒体基因组全序列为参考,对 Illumina 测序数据集 clean data 进行序列组装( $k=40,5$  次重复),获得的重叠序列群采用 99%可靠度输出一致性序列,使用 MITOS 网站<sup>[4]</sup>进行基因注释. 注释结果通过人工校正获得完整的苏姜猪线粒体全基因组信息.

1.3 线粒体基因组注释结果分析

得到正确完整的苏姜组线粒体基因组注释信息后,使用在线软件 tRNAscan-SE search server 以及 MITOS 网站得到 22 条 tRNA 的二级结构图.

1.4 线粒体基因组遗传多样性分析

从 GeneBank 中下载中国 6 个类型家猪的线粒体基因组全序列(表 1),采用 clustalX<sup>[5]</sup> 比对,采用 DNAsp6.12 软件<sup>[6]</sup>进行核苷酸多样性分析,去除高度变异的控制区后,采用 MEGA7.0<sup>[7]</sup>进行 NJ 法系统树构建,参数设置分别为:1 000 次自展检验,kimura 双参数模型,突变速率为通用速率,同时考虑转换颠换,将空位和缺失视为完全的删除位点.

表 1 不同类型家猪的线粒体全序列信息  
Table 1 Mitochondrial complete sequence information of different types of domestic pigs

类型	品种	GenBank 登录号	序列长度	类型	品种	GenBank 登录号	序列长度
江海型	梅山	JN601071.1	16 770	西南型	荣昌猪	KM044239.1	16 710
	姜曲海	AF486872.1	15 977	华北型	民猪	AF486864.1	15 977
	二花脸	AF486861.1	15 978	华中型	宁乡猪	KF472178.1	16 690
华南型	香猪	KC250273.1	16 709	高原型	藏猪	AF486856.1	15 977

## 2 结果与分析

由校正后的一致性序列可知,苏姜猪的线粒体基因组全长 15 436 bps,共有 37 条线粒体基因组基因,碱基组成 T:25.7%,C:26.3%,A:13.6%,G:34.5%,其线粒体基因组碱基组成 A+T 约占 40%,C+G 约占 60%,线粒体基因排列紧凑,一些相邻的基因之间发生了碱基对的重排现象,22 个 tRNA 中总共有 8 处错配,其中在 *trnI*~*trnQ* 基因之间重叠了 3 bp,在 *atp8*~*atp6* 基因之间重叠了 43 bp,在 *nad4L*~*nad4* 基因之间重叠 7 bp,在 *nad5*~*nad6* 基因之间重叠了 17 bp,在 *trnC*~*trnY*、*atp6*~*cox3* 基因之间重叠了 1 bp,其他相邻基因的连接以及连接顺序如表 2.

表 2 苏姜猪线粒体全基因组基因分布情况

Table 2 Sujiang pig's mitochondrial genome-wide gene distribution

基因名称	开始位置	终止位置	所在链	长度	基因间核苷酸数
<i>trnF</i> ( <i>ttc</i> )	1	70	+	70	
<i>rrnS</i>	71	1031	+	961	
<i>trnV</i> ( <i>gta</i> )	1032	1099	+	68	
<i>rrnL</i>	1100	2669	+	1570	
<i>trnL2</i> ( <i>tta</i> )	2670	2744	+	75	2
<i>nad1</i>	2747	3701	+	955	
<i>trnI</i> ( <i>atc</i> )	3702	3770	+	69	-3
<i>trnQ</i> ( <i>caa</i> )	3768	3840	-	73	1
<i>trnM</i> ( <i>atg</i> )	3842	3911	+	70	
<i>nad2</i>	3912	4953	+	1042	
<i>trnW</i> ( <i>tga</i> )	4954	5021	+	68	6
<i>trnA</i> ( <i>gca</i> )	5028	5095	-	68	1
<i>trnN</i> ( <i>aac</i> )	5097	5171	-	75	32
<i>trnC</i> ( <i>tgc</i> )	5204	5269	-	66	-1
<i>trnY</i> ( <i>tac</i> )	5269	5334	-	66	1
<i>cox1</i>	5336	6880	+	1545	3
<i>trnS2</i> ( <i>tca</i> )	6884	6954	-	71	5
<i>trnD</i> ( <i>gac</i> )	6960	7027	+	68	
<i>cox2</i>	7028	7715	+	688	
<i>trnK</i> ( <i>aaa</i> )	7716	7782	+	67	1
<i>atp8</i>	7784	7987	+	204	-43
<i>atp6</i>	7945	8625	+	681	-1
<i>cox3</i>	8625	9408	+	784	
<i>trnG</i> ( <i>gga</i> )	9409	9477	+	69	
<i>nad3</i>	9478	9823	+	346	1
<i>trnR</i> ( <i>cga</i> )	9825	9893	+	69	
<i>nad4L</i>	9894	10190	+	297	-7
<i>nad4</i>	10184	11561	+	1378	
<i>trnH</i> ( <i>cac</i> )	11562	11630	+	69	
<i>trnS1</i> ( <i>agc</i> )	11631	11689	+	59	
<i>trnL1</i> ( <i>cta</i> )	11690	11759	+	70	
<i>nad5</i>	11760	13580	+	1821	-17
<i>nad6</i>	13564	14091	-	528	
<i>trnE</i> ( <i>gaa</i> )	14092	14160	-	69	4
<i>cob</i>	14165	15304	+	1140	
<i>trnT</i> ( <i>aca</i> )	15305	15372	+	68	
<i>trnP</i> ( <i>cca</i> )	15373	15436	-	64	

13 条蛋白编码基因中有 1 条基因(*nad6*)在轻链编码,12 条基因(*nad1*、*nad2*、*nad3*、*nad4*、*nad4L*、*nad5*、*nad6*、*cox1*、*cox2*、*cox3*、*atp6*、*atp8*)在重链编码.

DNASP 计算 9 个线粒体全基因组得到的多态信息位点为 32 个,核苷酸多态性  $\pi=0.000\ 53$ ,计算 4 个江海型线粒体全基因组得到的多态信息位点为 18 个,核苷酸多态性  $\pi=0.000\ 57$ . 可见江海型猪线粒体基因多态性要高于全国 6 个家猪类型的平均水平.

由图 1 可知,苏姜猪和二花脸聚为一支,和梅山关系最远. 由图 2 可知,苏姜猪在数据量较大的情况下,单独分为一支,二花脸反而和华南型和西南型的猪聚在了一起,华北型和梅山聚在了一起,和其他猪种关系最远.

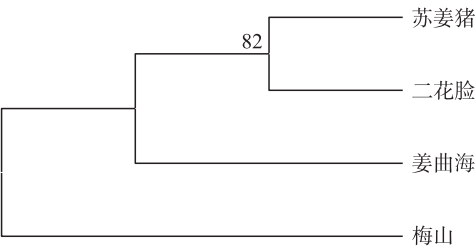


图 1 基于线粒体基因组的四种江海型猪系统发生树  
Fig. 1 Phylogenetic tree of four types of Jianghai porcine based on mitochondrial genome

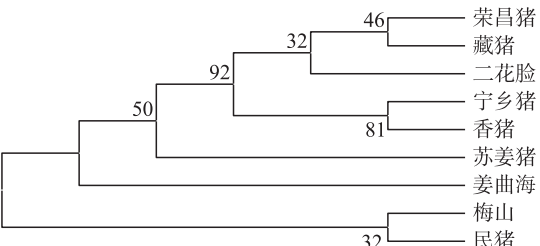


图 2 基于线粒体基因组的六种类型猪系统发生树  
Fig. 2 Six types of porcine phylogenetic trees based on mitochondrial genomes

22 条 tRNA 基因的长度和结构如图 3 所示. 大多数的 tRNA 的二级结构能形成三叶草结构. 对比家猪类线粒体基因组原始的基因排序,苏姜猪线粒体基因组变异不大,其线粒体全基因组的基因排列顺序为: *trnF-rrns-trnV-rrnL-trnL2-trnI-trnQ-trnM-nad2-trnW-trnA-trnN-trnC-trnY-cox1-trnS2-trnD-cox2-trnK-*

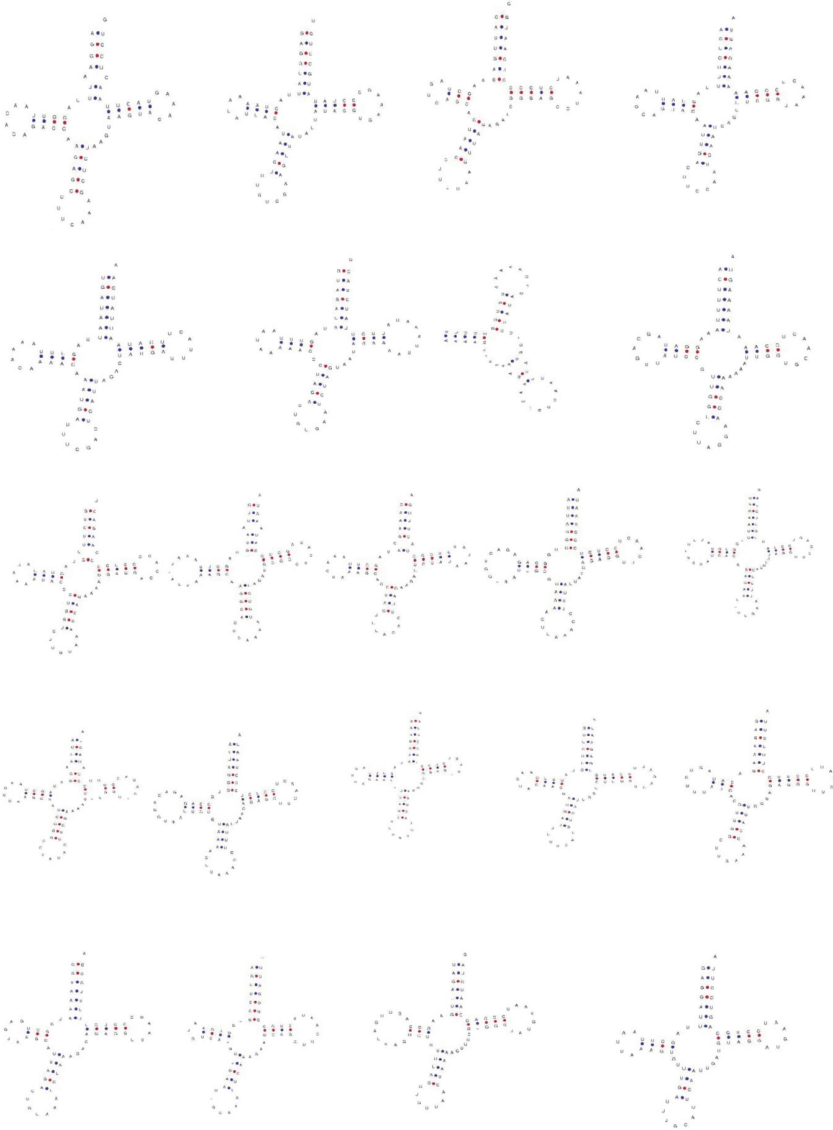
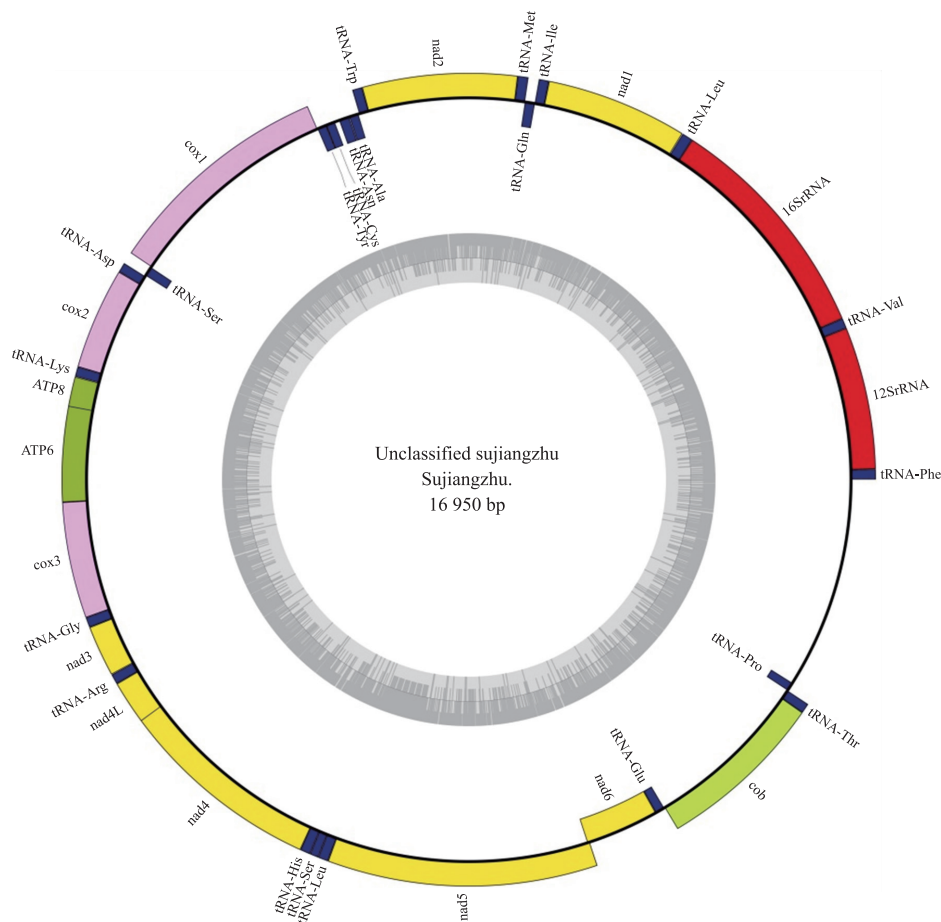


图 3 苏姜猪 22 条 tRNA 二级结构  
Fig. 3 The secondary structure of Sujiang pig's tRNA



- 
- [2] 《中国家畜家禽品种志》编委会/《中国猪品种志》编写组. 中国猪品种志[M]. 上海:上海科学技术出版社,1986.
- [3] DRUMMOND A J, ASHTON B, CHEUNG M, et al. Geneious v4.8.5, <http://www.geneious.com>.
- [4] BERNT M, DONATH A, JÜHLING F, et al. MITOS: Improved de novo metazoan mitochondrial genome annotation[J]. *Molecular phylogenetics and evolution*, 2013, 69(2): 313–319.
- [5] LARKIN M A, BLACKSHIELDS G, BROWN N P, et al. Clustal W and Clustal X version 2.0[J]. *Bioinformatics*, 2007, 23: 2947–2948.
- [6] ROZAS J, FERRER M A, SÁNCHEZ D, et al. DnaSP 6: DNA sequence polymorphism analysis of large datasets[J]. *Mol Biol Evol*, 2017, 34: 3299–3302.
- [7] KUMAR S, STECHER G, TAMURA K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets[J]. *Molecular biology & evolution*, 2016, 33(7): 1870.
- [8] 张亚平, 施立明. 动物线粒体 DNA 多态性的研究概况[J]. *动物学研究*, 1992, 13(3): 289–298.
- [9] 黄勇富, 张亚平, 邱祥聘, 等. 猪线粒体 DNA 多态性与中国地方猪种起源分化的关系[J]. *Journal of genetics & genomics*, 1998, 25(4): 322–329.
- [10] 胡慧艳, 贾青, 刘津, 等. 中国地方猪种线粒体基因组遗传多样性及其系统发育分析[J]. *中国畜牧兽医*, 2017, 44(8): 2360–2368.

[责任编辑:黄 敏]