

# 醋栗番茄的基因组为番茄育种过程中的结构变异提供了新的见解

宫亚运

贝瑞基因科技服务部

# 文章概括

---

## 醋栗番茄的基因组为番茄育种过程中的结构变异提供了新的见解

Genome of *Solanum pimpinellifolium* provides insights into structural variants during tomato breeding

- ◆ 关键词语：醋栗番茄，基因组，SV，eQTL
- ◆ 研究单位：美国博伊斯汤普森植物研究所
- ◆ 发表期刊：Nature Communications
- ◆ 发表时间：2020.11.16
- ◆ 影响因子：12.121

# 思路设计

- ✓ 1、醋栗番茄的高质量基因组组装
- ✓ 2、醋栗番茄与近缘物种的比较基因组分析
- ✓ 3、醋栗番茄与现代栽培品种Heinz 1706变异检测，重点检测SV事件
- ✓ 4、番茄群体的基因型分型
- ✓ 5、番茄驯化育种中SVs的筛选
- ✓ 6、与育种性状相关的SV选择
- ✓ 7、eQTL分析
- ✓ 8、联合GWAS分析进行性状关联
- ✓ 9、候选基因的验证工作

# 物种背景

番茄原本生长在南美洲的智利秘鲁一带的安第斯山脉中,是小小的醋栗番茄.16世纪的时候,西班牙殖民者把这些可爱的小东西带回了欧洲.



醋栗番茄为茄科，番茄属。是普通番茄种野生型番茄亚种中的一个变种。茎蔓生，多分枝，表面光滑无绒毛，无特殊的番茄气味。叶小，裂片圆形或心脏形。花序长20~25 cm，最长40 cm，每序着花30~40朵。果实小、圆球形，红色或黄色。番茄品质和抗病育种的重要原始材料。果实中干物质含量高达8%~10%，植株对番茄青枯病、叶霉病等多种病害有很强的抗性。

# 材料方法



## 材料情况

醋栗番茄LA2093在纽约伊萨卡博伊斯·汤普森研究所的温室中生长，光照时间为16小时，白天为25°C，夜间为20°C



## 二代测序

Hi-C，构建450bp的文库，Illumina NextSeq 500 平台



## 三代测序

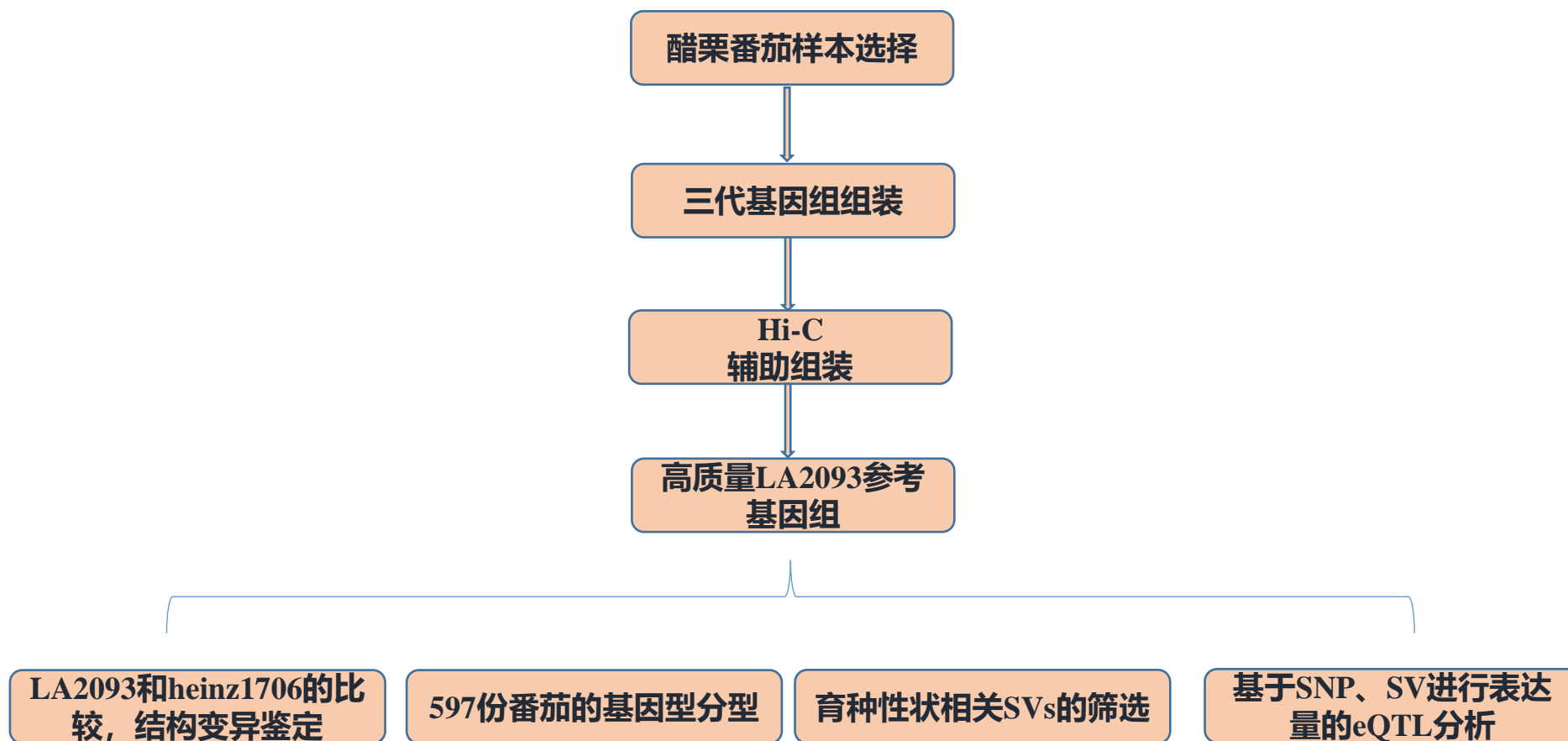
CLR，103X，PacBio Sequel 平台



## RNA测序

从未成熟、成熟的绿色和红色成熟果实中提取总RNA，对每个样本进行三次生物重复，在Illumina HiSeq 2000平台上测序

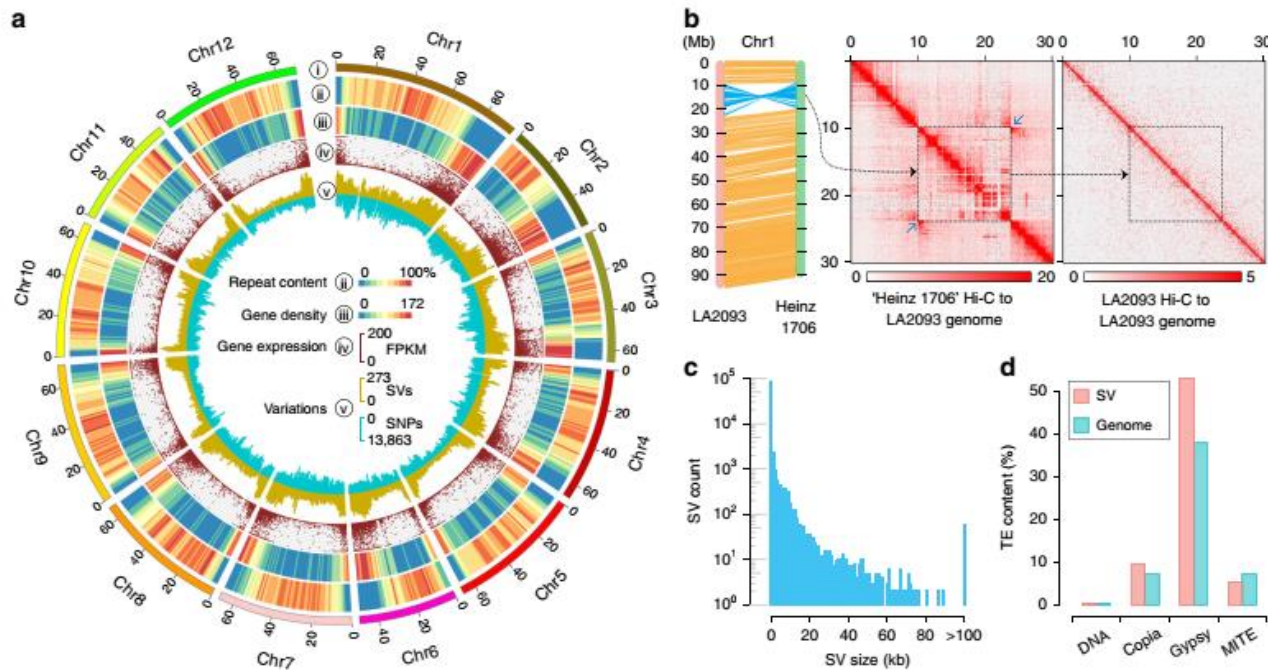
# 技术路线



本次研究的技术路线



# 研究结果1——醋栗番茄的基因组组装和两个番茄间的SVs



醋栗番茄基因组以及醋栗番茄LA2093与现代栽培种番茄Heinz 1706之间的结构变异鉴定

1、PacBio+Hi-C数据，组装基因组大小为807Mb，Contig N50为10.9Mb，BUSCO完整性为97.8%；远高于之前组装版本。

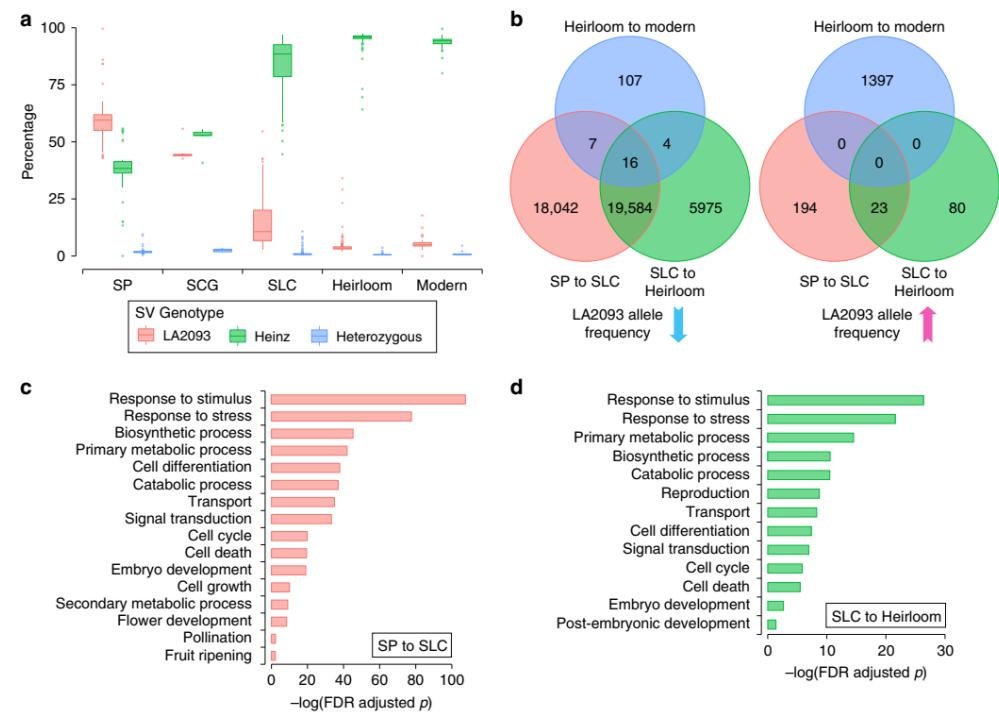
2、LA2093和heinz1706的基因组序列比对显示两个参考基因组之间具有良好的共线性。尽管存在高共线性，共有28个从483bp到13.9mb的倒位分布在所有12条染色体上。这些倒位总共包含大约800个基因，其中很大一部分（457个）被注释为与抗病性和对非生物胁迫的反应相关的功能。

3、两个基因组间鉴定出92523个InDels，范围从10bp到2.4mb，大多数InDels 相对较短，82.8%<100bp，只有0.1%>100kb。

4、研究发现InDels富含刺激、生殖、信号转导、初级和次级代谢过程的基因，这些InDel可能是野生番茄与栽培番茄抗病性和果实品质性状差异的原因之一。

5、检测到了几个与番茄驯化性状有关的SVs，如CSR基因中1.4kb的缺失和22bp的插入导致果实重量增加，LNK2基因中7kb的缺失导致番茄昼夜节律延长，TomLoxC基因启动子中的4kb替换，有助于水果风味，以及调控花分生组织活性的ENO基因启动子中的85 bp缺失。

# 研究结果2——番茄驯化育种中SV的选择

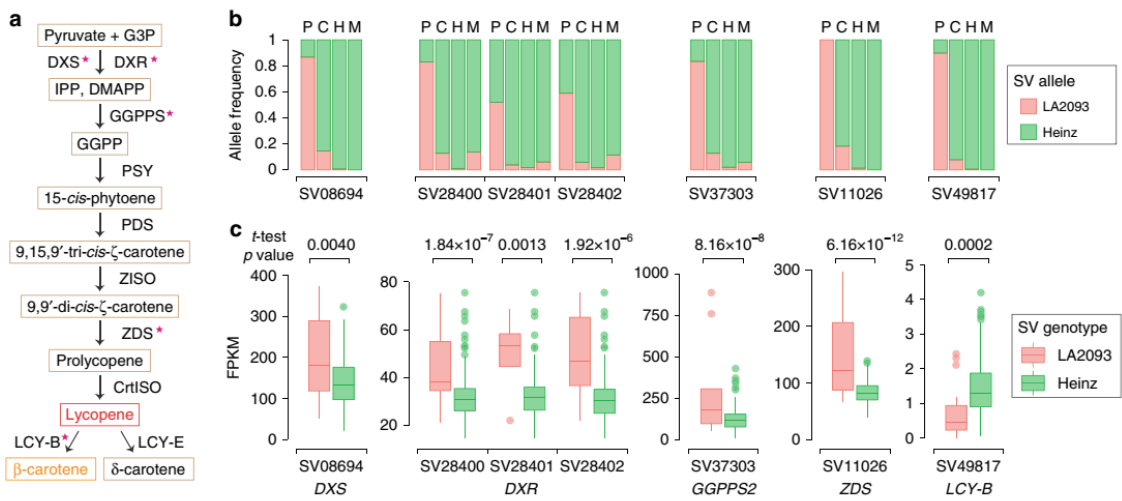


番茄驯化育种中SV的选择

- 1、对597份番茄进行分析，包括51份SP、6份SCG、228份SLC、226份古老品种、52份现代品种和34份其他品种。系统发育和种群结构分析清楚地将SP和SCG群从古老群体和现代类群中分离出来，SLC是介于野生和栽培材料之间的中间类群。
- 2、不同番茄群体间SV等位基因频率的变化是进化过程的结果，如选择理想性状、减少群体规模和从祖先群体中导入。为了鉴定番茄驯化育种过程中选择的SV，文章研究了SV等位基因频率从SP到SLC进行驯化，从SLC到SLL古老品种的变化，以及从古老品种到现代优良品系的频率变化（图a，b）。
- 3、在驯化和改良过程中，受SVs影响的基因之间共享丰富的功能，包括胁迫和刺激反应、生物合成、细胞分化、胚胎发育、授粉和生殖过程（图c，d）。这些结果表明，在番茄驯化和改良中，heinz1706等位基因具有共同的选择偏好。尽管野生物种的等位基因从SP持续丢失到SLC和SLL，1397个SVs在现代品系中表现出明显高于古老品种的LA2093等位基因频率（图2b）。这可能与野生种的农艺有利等位基因重新引入现代SLL品系有关。
- 4、结果表明，现代品种的SP导入可能是通过当代育种获得的，从而引入非生物和生物胁迫耐受性以及良好的风味性状。



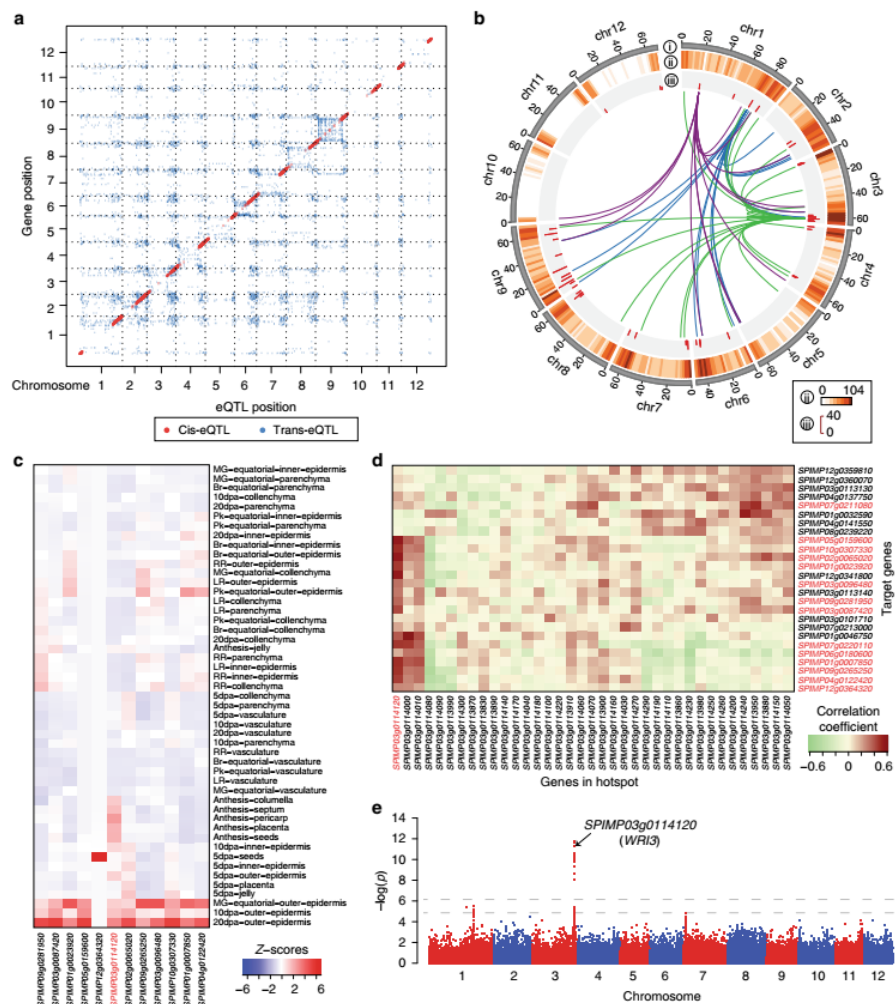
# 研究结果3——与育种性状相关的SV的选择



选择影响番茄红素代谢基因表达的SVs

- 1、番茄果实性状在驯化和改良过程中发生了显著变化。研究SVs可能影响重要园艺性状基因表达或功能的种群动态，有助于我们更好地理解人类选择对这些基因的影响，为育种提供潜在的靶点。果实膨大是番茄主要的驯化综合征。一些SVs已被证实与果实重量有关，包括CSR的1.4kb缺失和ENO启动子的85bp缺失，这两个都是导致果实较大的原因。
- 2、发现heinz1706等位基因频率从SP到SLC显著增加。这些Heinz 1706等位基因，大多在启动子中缺失，这些突变可能通过下调番茄红素生物合成基因而导致SLC和SLL中番茄红素水平的降低，并通过上调LCY-B促进番茄红素降解。（左图）
- 3、番茄FUL1、FUL2和TAGL1与RIN相互作用，在果实成熟过程中起关键作用，研究在 heinz1706 的 FUL1 的 CDS 中发现了一个 15bp 的框内缺失（SV28464），结果表明，FUL1缺失等位基因与FUL1的高表达相关，并在驯化过程中被积极选择。
- 4、将野生种等位基因导入SLL已成为番茄育种中提高抗病性的一种策略。我们在许多研究良好的抗病基因中鉴定了选定的SVs。对这些SVs的进一步功能鉴定有助于恢复栽培番茄的抗病特性。

# 研究结果4——eQTL分析



全基因组范围内的eQTL

1、为了探索SV在基因表达调控中的作用，研究者使用公开的转录组数据进行了eQTL分析。已鉴定出在10,789个eQTL区域中共有46,848个SV与5595个基因的表达显著相关，包括2708个（25.1%）顺式和8081个（74.9%）反式eQTL。顺式-eQTLs与基因表达更显著相关，因此比反式-eQTLs解释了更多的表达变异。大多数顺式eQTL的前导SV位于编码区的开始或末端附近，表明这些区域中序列对基因表达的重要调节作用。

2、总共确定了48个调控554个基因表达的热点以及每个热点中潜在的主调控子。在3号染色体上发现了一个主要热点，具有26个靶基因，其中14个与脂质代谢有关。番茄WRI3与脂肪代谢基因在发育中果实表皮的共表达表明，它在组织特异性调节脂类生物合成中的潜在功能，从而有助于果实角质层蜡的积累（图4c）。WRI3的表达与其所有靶基因的表达呈正相关（图4d），表明其正调控作用。此外，eQTL分析确定了两个SVs，第六内含子的11bp插入和WRI3启动子的77bp缺失，它们与表达显著相关。这些结果表明WRI3可能是果实外表皮脂质沉积的关键调节因子。

# 文章亮点和不足

## 亮点

01

对于整个番茄品系的SV检测和筛选做的比较全面

02

找到了许多与番茄重要农艺性状相关的结构变异和基因，为后续番茄的育种工作带来了更多的启示和参考

## 不足

01

三代测序部分还是用的CLR模式，没有加上最新的HiFi测序和HiFi组装技术

02

没有结合基于传统表型的GWAS进行联合分析，后续也没有任何的验证试验，文章深度不够



# Thank You!



官方网站



官方微信