



基因组育种技术利用及其优势

北京康普森农业科技有限公司

汇报人：冯羿方

时间：2019.4.21



康普森生物公司简介

北京康普森生物技术有限公司是专注于中国农业基因组产业化发展，提供创新型农业基因组检测技术和育种整体解决方案的国家级高新技术企业。公司具有基因芯片检测平台和生物信息学分析平台，主要服务内容为个性化育种方案制定、畜禽基因组选育、选配策略制定、遗传疾病评估、功能基因检测等。通过一站式育种服务，帮助育种工作者更高效地进行选育与选配工作。公司由4个子公司组成：农业产业公司、检测公司、华中和西南技术服务子公司和全国近10个销售分支机构，技术服务辐射全国。



全国畜牧基因组产业转化高峰论坛

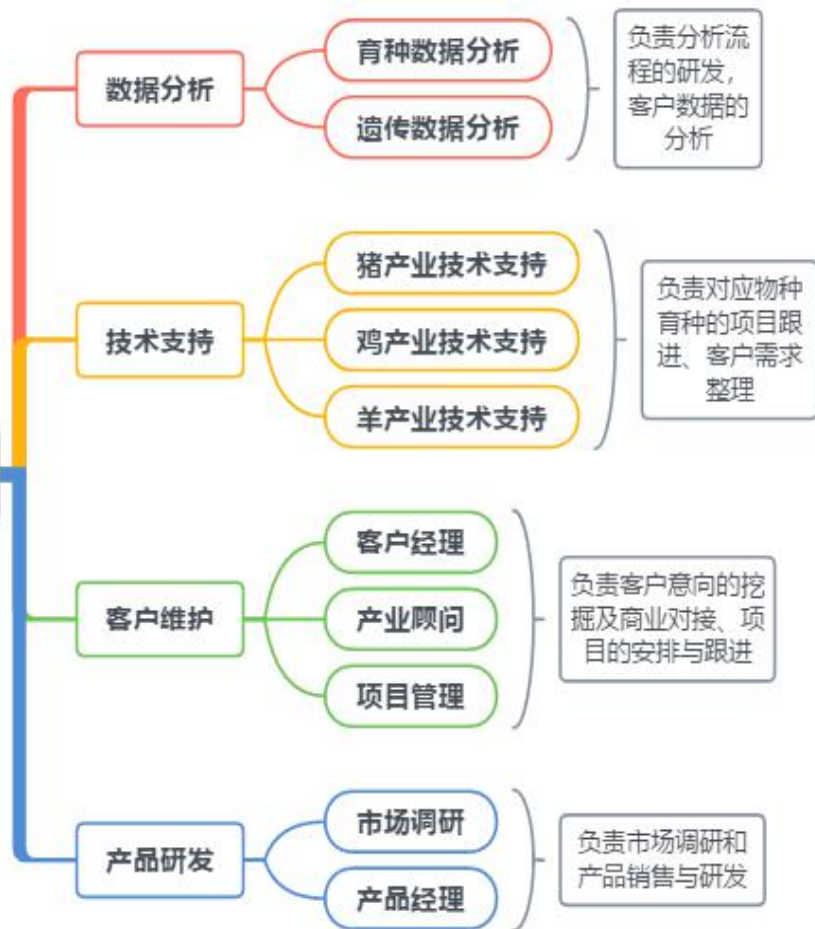
为了响应国家全面实施畜牧种业遗传改良计划的号召，进一步推广基因组选择技术和产业化应用，康普森生物从2015年起每年都会主办“全国畜牧基因组产业转化高峰论坛”。邀请中国农业大学、中国农业科学院北京畜牧兽医研究所、南京农业大学、四川铁骑力士牧业科技有限公司、广东温氏集团等单位的多位畜牧遗传学领域的专家共同分享科研成果，探讨行业热点，共同推动民族品牌的畜牧育种研究和技术革新！



北京康普森农业



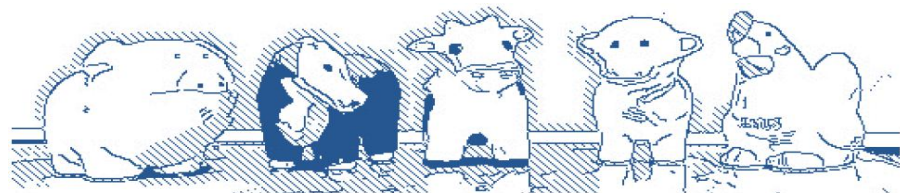
北京康普森农业团队



北京康普森农业科技有限公司
为北京康普森生物技术有限公司的全资子公司，
是康普森生物最重要的业务板块。

- 致力于农业基因组科研成果转化，以基因组育种的产业化应用服务于中国畜牧产业的发展。
- 搭建有基因组育种服务平台，通过技术研发提高了基因组检测技术在国内的应用。
- 与畜禽产业各方等合作，共同打造联合育种产学研院企合作平台。

实现“专注分子育种，用基因科技解决选育问题”的企业愿景。





目录 Catalog



01



技术的发展及时代背景

基因组选择的主要流程及技术难点



02

03



芯片技术在种质资源方面的应用



01

育种的发展及基因组选择

为什么要使用基因组选择？

育种工作的价值

针对与消费者

满足社会

优质、安全
廉价、美味

肉质性状（IMF、肉色、嫩度、PH等）
胴体性状（瘦肉率、背膘）
健康度高（减少疫苗、兽药使用）

养猪产业

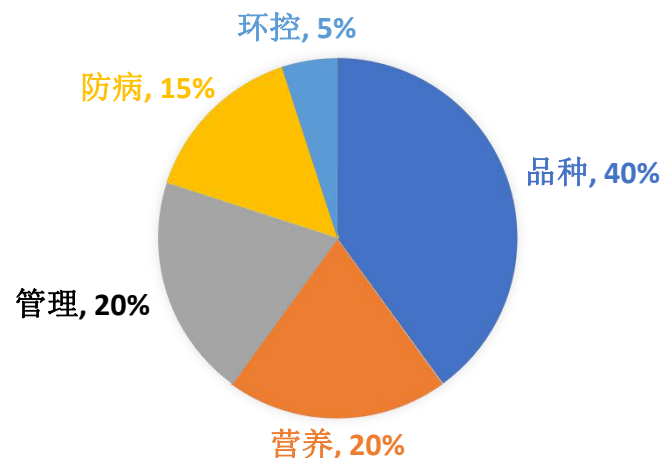
针对养猪企业

提升效益

造肉成本低
规模化生产

商品猪：活得多、长得快、屠宰率高、FCR高
繁殖母猪：生得多、使用年限长 高度集约化条件下的生物安全

产业发展贡献率



育种工作的流程



**制定方向，坚持选育，实现
群体经济性状不断提升。**

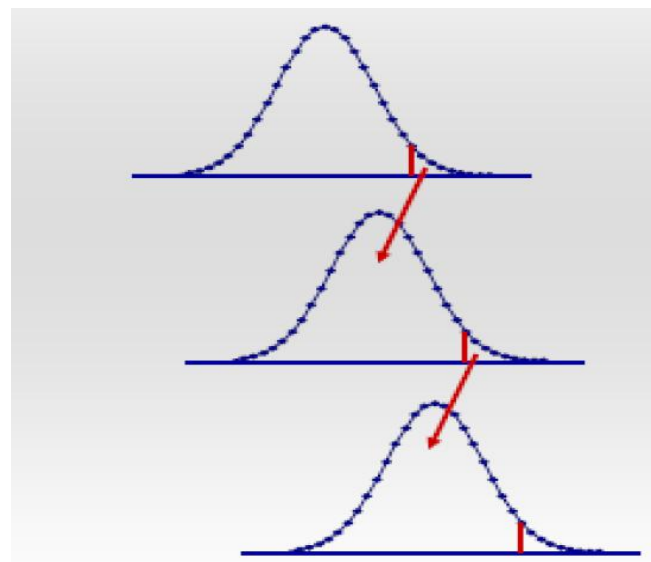
遗传进展：

$$\Delta G_t = \frac{i \cdot \sigma_A \cdot r_{AI}}{T}$$

遗传变异性(σ)：这是遗传改良基础，取决于**选育的群体**

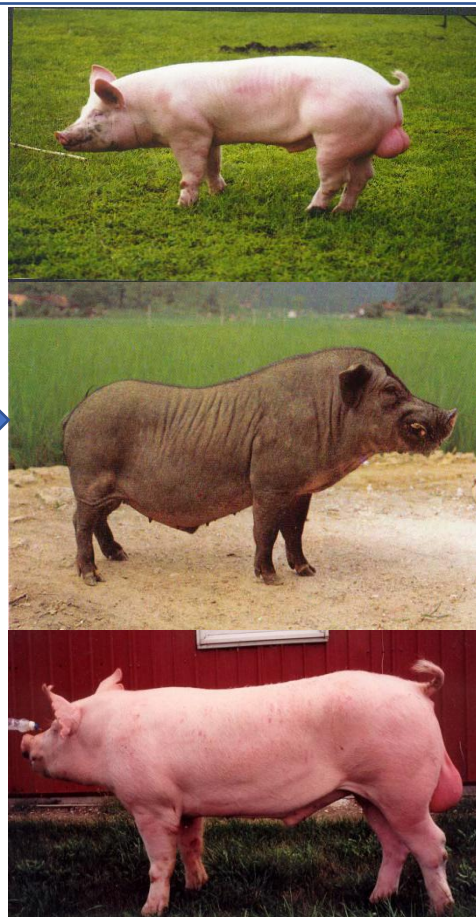
选择强度(i)、世代间隔(T)：依据**育种措施和育种方案**

选择的准确性(r)：取决于**测定的准确性**和**育种值评估的方法**



育种值的定义

数量性状表型值： **表型值** (P) = 遗传效应 (G) + 环境效应 (E)
= 加性遗传效应 (A) + 显性效应 (D) + 上位效应 (I)
+ 系统环境效应 (ES) + 随机环境效应 (ER)



育种值估计的发展



育种值估计的**初级**阶段

选择指数法

根据目标性状
的表型值

经典育种时期

目前广为应用的

BLUP法

根据家系及系谱
资料（加性效应）

育种值估计方法

MBLUP法

根据经济性状的
遗传标记

分子育种时期

最新的育种值估计方法

Genomic
Selection

表型值+基
因组效应值



BLUP产生背景

◆ 应用选择指数法的前提

- ◆ 不存在系统环境效应，或已校正；
- ◆ 被估个体与表型观测值来源于同一总体；
- ◆ 遗传参数已估计出来；

满足上述条件时，选择指数的成果：例如鸡、猪的场内遗传评估中，利用选择指数获得较好效果；

这些条件难以满足时，选择指数的应用受到制约

❖ BLUP = 最佳线性无偏预测



$$y = Xb + Zu + e$$

y : 观测值向量，如：母牛产奶量、猪日增重

b : 固定效应向量，如：牧场、年季、胎次

u : 随机效应向量，如：个体育种值向量

e : 随机残差向量

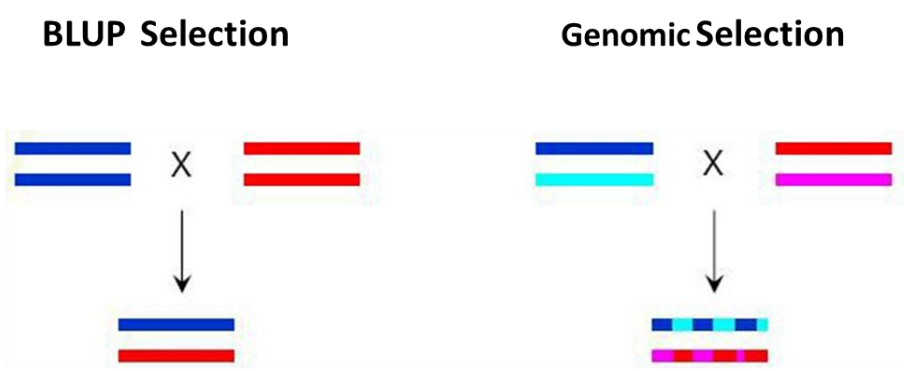
X 、 Z : 分别是与固定效应和加性遗传效应对应的关联矩阵

前提：需要大量测定的数据和三代以上的系谱资料

常规育种方法局限

Traits性状	Heritability (h ²)遗传力
Number farrowed 产仔数	0.10
Number weaned 断奶数	0.10
Birth weight 出生重	0.20
Weaning weight 断奶重	0.20
Feed efficiency 饲料报酬	0.25
Growth rate 生长速度	0.30
Age at puberty 初情期	0.35
Backfat thickness 背膘厚	0.40
Yield of lean cuts 瘦肉率	0.40
Loin eye area 眼肌面积	0.45

- ◆ 低遗传力性状效果不明显（比如繁殖力，母猪使用寿命）
- ◆ 测定成本高或无法测定性状（抗性、肉质等）
- ◆ 早期选择效果差（仅仅依靠系谱信息和亲属表型）
- ◆ 无法同时权衡多个性状的育种要求

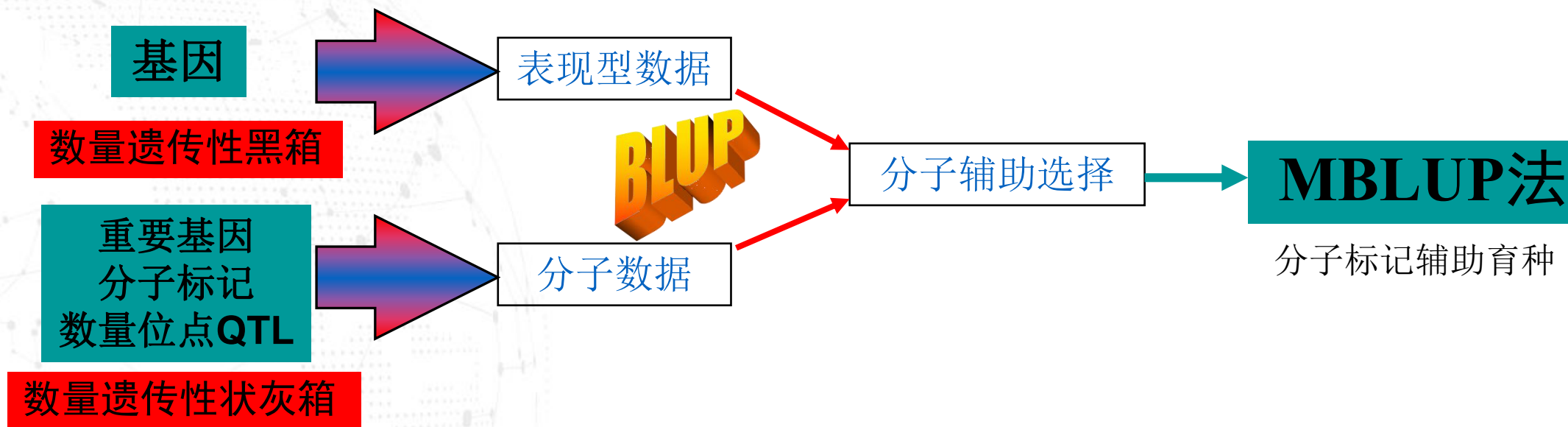


后代EBV = 父母均值
选择准确性 ~40%

后代EBV = 父母均值+DNA样本
选择准确性 ~60%

MBlup分子标记辅助选择

分子育种——将分子生物学技术应用于育种中，在分子水平上进行育种。
通常包括：分子标记辅助育种和遗传修饰育种（转基因育种）。



利用表型数据和分子数据的组合，减少试验数量，提高育种效率

分子标记应用的局限及全基因组选择的产生

分子标记应用的局限

- ❖ 分子标记的数量过少
- ❖ 分子标记只能解释非常少的遗传方差分量
- ❖ 基因分型的成本过高
- ❖ 分子标记/数量性状位点（QTL）的贡献量估计值不稳定
- ❖ 贡献量大的位点倾向于被高估
- ❖ 分子标记贡献量之在家系和控制试验中有意义
- ❖ 与遗传背景 和（或者） 环境的交互作用显著
- ❖ 分子标记与QTL的连锁不平衡只存在于独立群体

解决方案：全基因组选择

- ❖ 利用全部的SNPs
- ❖ 同时估计50000个SNPs的贡献量
- ❖ 解大型线形方程组

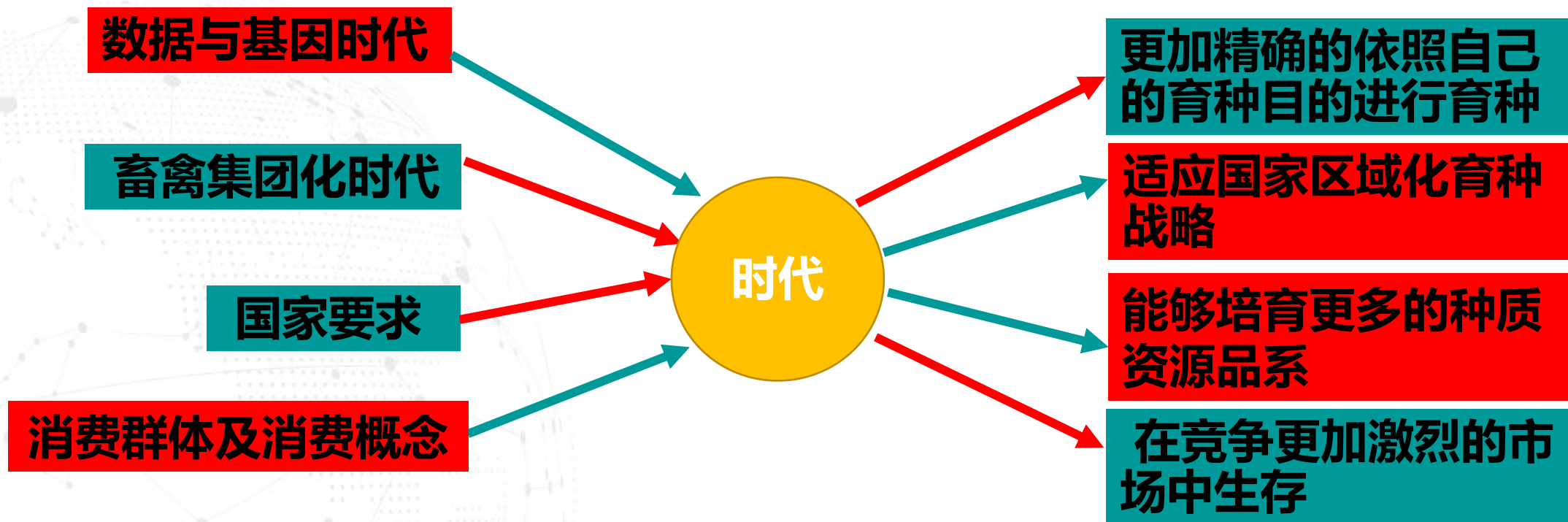
$$y_i = \mu + \sum_{\text{SNP } k} \beta_k g_{ik} + e_i$$

SNP effect
'1' alleles (0/1/2)

分子标记贡献量的估计值

利用分子标记
贡献量的估计
值估计其他基
因型的表现型

当代育种的需求和理念的变化



国外生猪的GS应用成果

➤ 采用先进的分子技术进行育种——PIC

Genomic Selection 1.0

采用**芯片技术**进行GS选育，10年开始进行，13年已经全面进行选育至今；

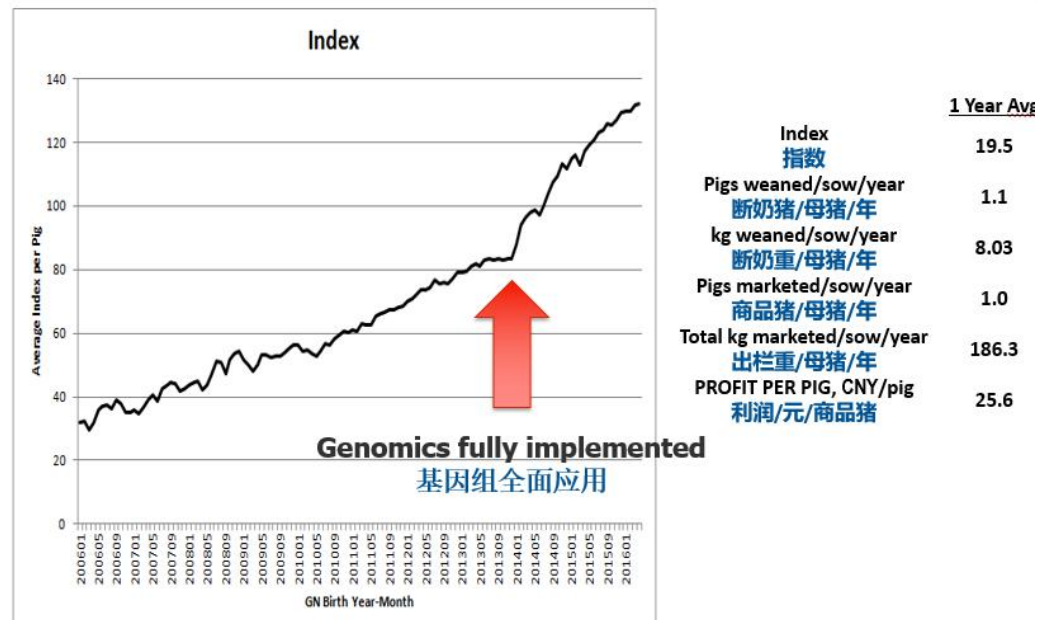
采用的是一步法GS，效果十分明显

Genomic Selection 2.0

利用高通量测序技术、人工智能AI、芯片技术等进行猪场育种。从2013年开始，做了大量的GWAS、克隆及基因编辑等实验，试图找出主效QTN，为下一个时代进行技术储备

育种是个需要检测且持续的过程，需要大量的投入，但是回报也是动人的

育种方面的研究投入大约为**3亿美金**



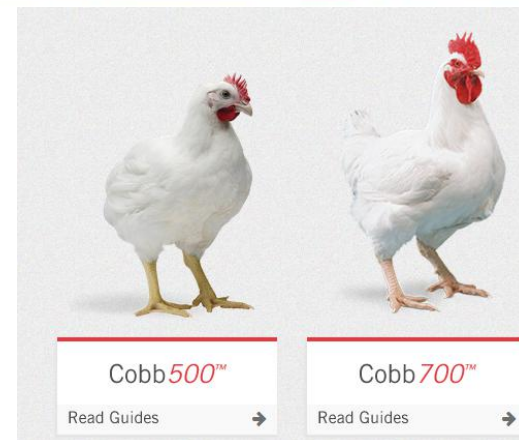
国外家禽产业的GS应用



美国科宝公司:

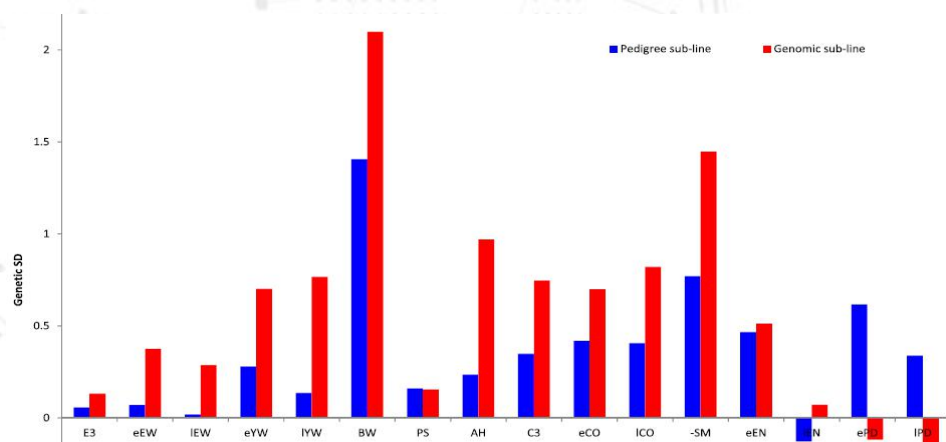
- ◆ 2008年开始与Hendrix、USDA合作5年，共投入研发资金1千万美元；2012年又与Hendrix签订了持续合作的备忘录
- ◆ 与常规选育比较，研究GS方法在育种群体重、胸肌重及腿病选育提高中的应用效果
- ◆ GS技术全面纳入常规育种体系

参考群>1.5万

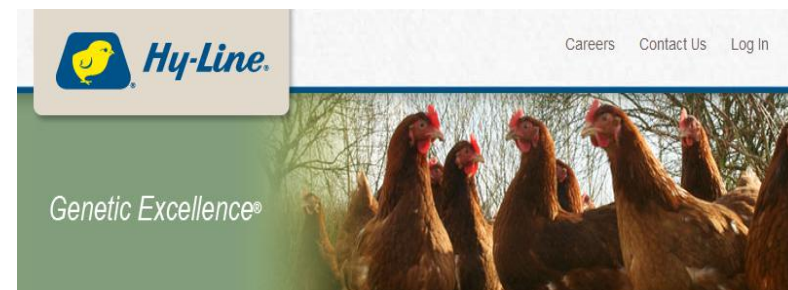


海兰公司GS选育效果-3年

- ◆ 受选择的16个性状中，蛋重、蛋黄重、蛋壳颜色、产蛋量等14个性状的GS选择效果均优于传统家系选择。



GS
传统



中国荷斯坦牛GS项目



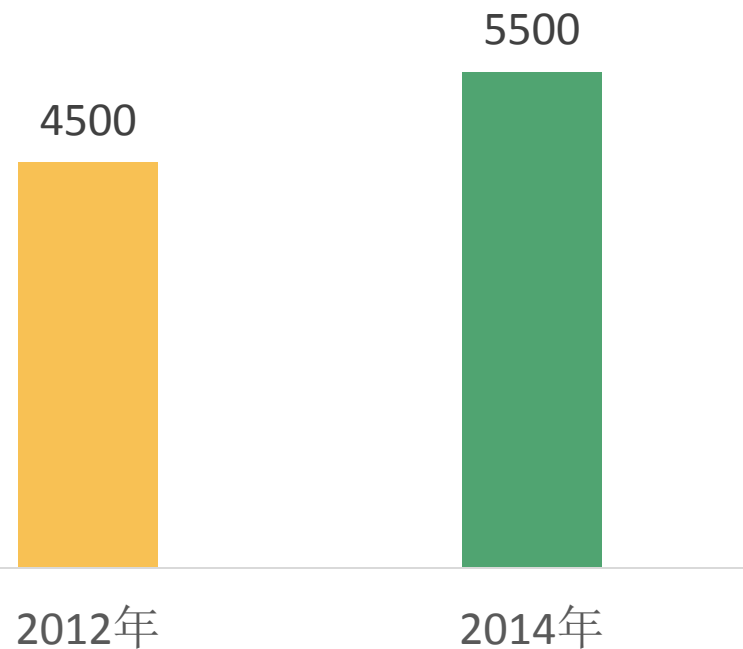
项目背景：

- 2012年，在农业部支持下，由中国农大承担，北京康普森参与，中国荷斯坦牛GS技术平台的研发；
- 建立参考群体（7183头母牛+273公牛），利用SNP芯片技术，对全国26个公牛站中的1547头进行遗传评估，选出724头优秀青年公牛；
- 该项目整体提升了我国奶牛育种水平，我国荷斯坦牛平均年产奶量从2010年的4500kg提高到2014年的5500kg；经济效益达21.87亿元；

项目奖项：

- 该技术体系2012年被农业部指定为我国荷斯坦青年公牛遗传评估的唯一方法；
- 中国荷斯坦牛基因组选择分子育种技术体系的建立与应用获得**2016年国家科技进步二等奖**；

中国荷斯坦牛平均年产奶量



我国首例采用全基因组选择育种的特级公猪在粤诞生



您现在的位置: 首页 > 全国

我国首例采用全基因组选择技术选育特级种公猪在粤诞生

日期: 2013-11-12 17:17 作者: 来源: 广东农业信息网

南方日报讯 昨日从华南农业大学获悉,通过3年多的技术攻关,我国首例采用全基因组选择技术选育的特级种公猪在广东温氏食品集团诞生。该项成果是我国在种猪育种技术上的重大突破,目前只有美国、丹麦等动物育种技术先进国家掌握了种猪全基因组选育技术。

据项目负责人吴珍芳教授、方美英教授介绍,该项研究是华南农业大学、广东温氏食品集团组建的国家生猪种业工程技术研究中心联合中国农业大学吴常信院士、李宁院士课题组和美国明尼苏达大学杨达教授课题组共同完成。研究人员采用高密度SNP芯片对杜洛克种猪进行全基因组扫描,获得覆盖整个猪基因组的标记信息,通过与对应的重要经济性状的表型信息进行全基因组关联分析,估计出染色体片段上每个标记的效应,最后建立基因组最佳线性无偏估测(GBLUP)方法。建立该方法后,采用全基因组选择可实现对候选个体从表型选择到基因选择的突破,解决动物个体肉质和抗性等性状难以选育的技术障碍,还可实现低成本的早期选择。



首例采用全基因组选择技术选育的特级种公猪,已经配种应用。

北京油鸡GS选育



北京油鸡 (50:250)



G3世代选择效果

组别	全基因组选择	随机留种对照群
N	131	99
胸肌 IMF%	2.93 ± 0.97^a	2.66 ± 0.88^b

G4 世代选择效果

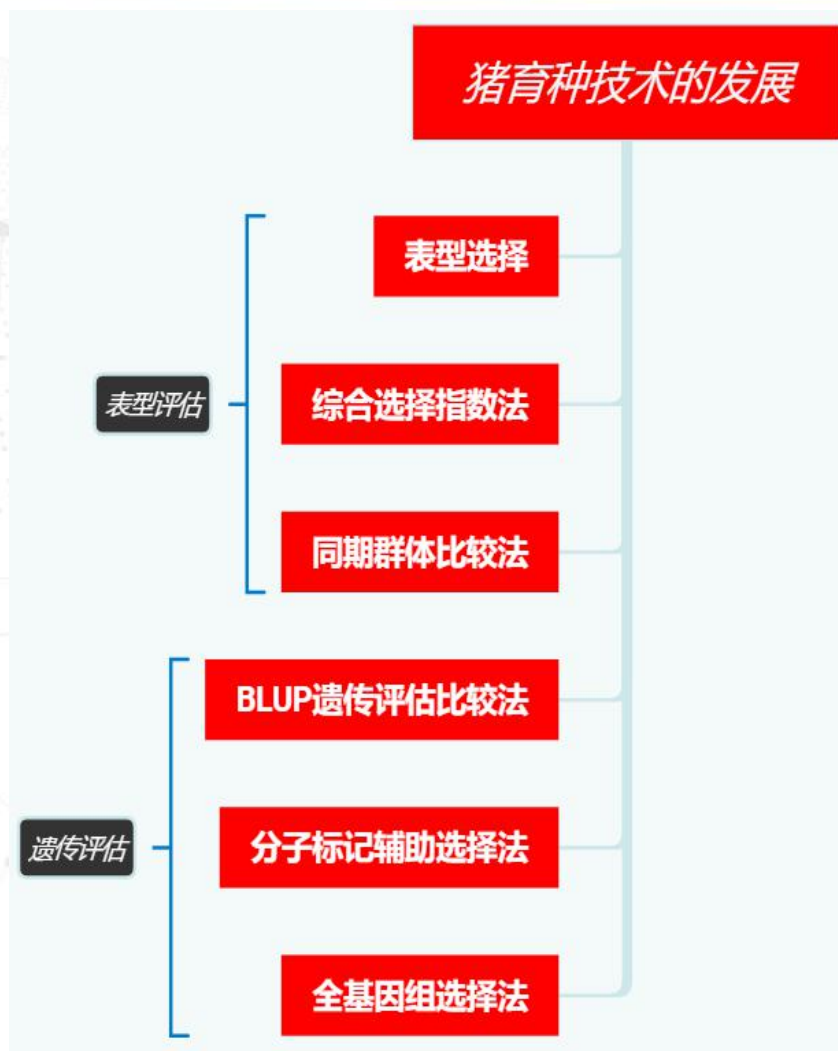
组别	全基因组选择	随机留种对照群
N	80	77
嫩度	3614.3^b	3697.1^c
胸肌 IMF%	2.15 ± 0.98^a	1.95 ± 0.91^b

提高**10%**!

基因组选择的发展



猪育种技术的发展



GS发展的回顾

2001年, Meuwissen等首次提出GS选择

2006年, Schafferr重新指出GS在奶牛育种中的巨大价值

2009年, GS技术在奶牛的选育过程全面使用; 产奶产奶性状遗传进展提高50%-100%, 女儿妊娠率、生产寿命和体细胞评分遗传进展提高300%-400%

2010年, 丹育 (DanAvl) 在其育种核心群对D、L、Y三大猪群实施GS, 每年测定10,000头个体基因型, 年遗传进展提高25%

2011年, 托佩克公司 (Topigs) 对公猪实施GS, 2012年对母猪实施GS, 使遗传进展提高30%

2012年, PIC公司在纯系育种中尝试实施GS; 2013年, 对所有核心群和所有性状均实现基因组评估, 采用SSGBLUP法选育, 遗传进展提高40%

A decorative background on the left side of the slide. It features a large, faint network of nodes and lines. Overlaid on this is a large red hexagon with a white border containing the number '02'. In the top left corner, there are two smaller hexagons, one grey and one white, both with 3D effects.

02

基因组选择的简介

A thin horizontal line with diamond-shaped endpoints at each end, positioned below the main title.

基因组选择的原理及技术细节

猪全基因组选择及其优势

基因组选择 (Genomic Selection, GS) 就是利用**基因组的遗传标记**信息计算育种值用于选种的方法。它就是全基因组范围上的标记辅助选择，是传统MAS选择的升级。

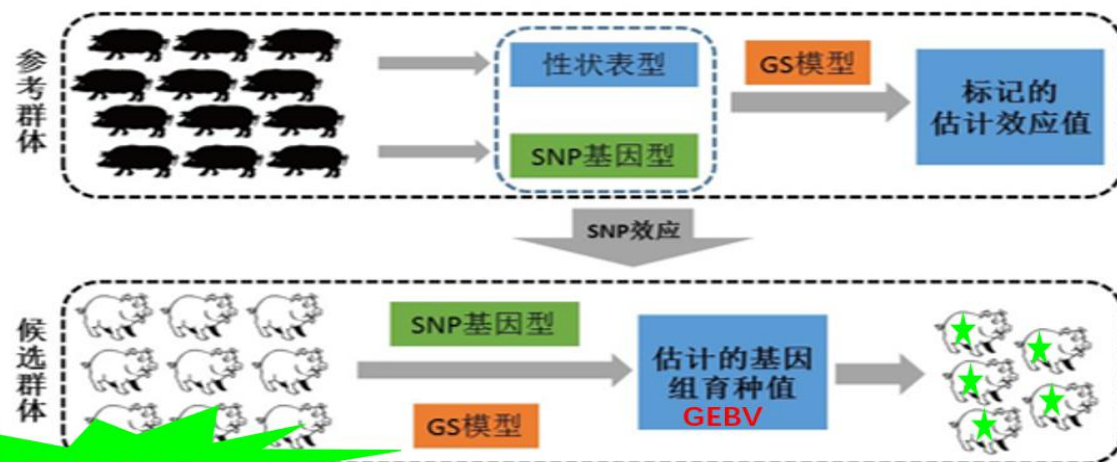
遗传进展:

$$\Delta G_t = \frac{i \cdot \sigma_A \cdot r_{AI}}{T}$$

特点：

- 1 通过芯片可以捕获大量的遗传变异，**提高选择的准确率**；
- 2 可对低遗传力和难以度量的性状进行选择，**降低测定成本**；
- 3 降低群体的近交增量，**防止近交衰退**。
- 4 可以实现早期选育，**减少企业后期的投入**。

技术的主要流程



基因组选择分析流程图

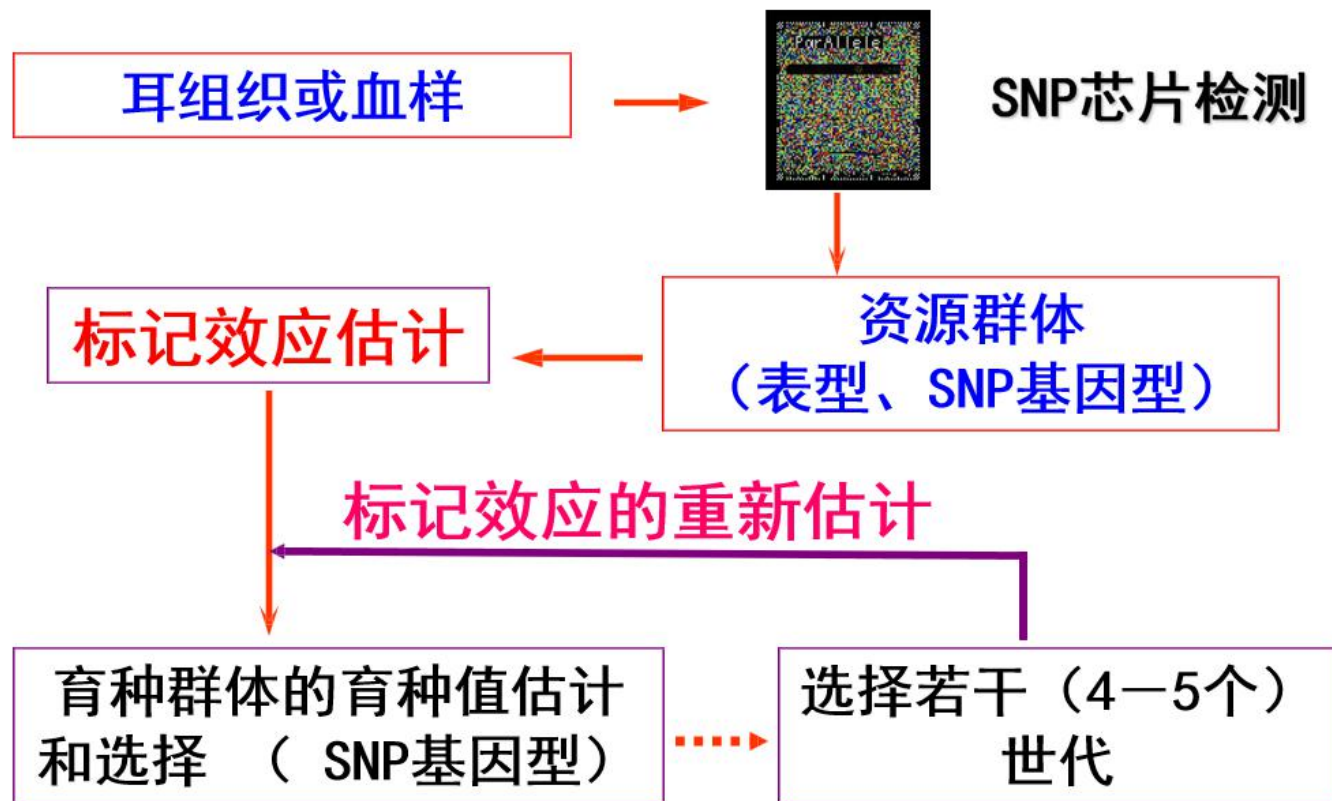
GS优势-提高选择准确性



遗传力	方法	选择准确性	绝对准确性	相对准确性 (%)
0.1	blup	0.22	0.101	45.91%
	ssblup	0.322		
0.3	blup	0.384	0.112	29.26%
	ssblup	0.497		
0.5	blup	0.449	0.052	11.57%
	ssblup	0.502		

因此，GS技术对于**肉质**、**繁殖**等低遗传力性状或者常规选育难以测定的性状选育效果更显著。

基因组选择的技术环节及挑战



获取个体基因组变异的成本及数据质量是这个技术应用的主要限制
高密度芯片是这个GS分型技术的主流

技术应用的难点

1 适合育种使用的基因分型技术

- 高质量位点且均匀分布
- 要含有大量与选育性状相关联的主效标记
- 要覆盖现今生产主要使用的品种或品系

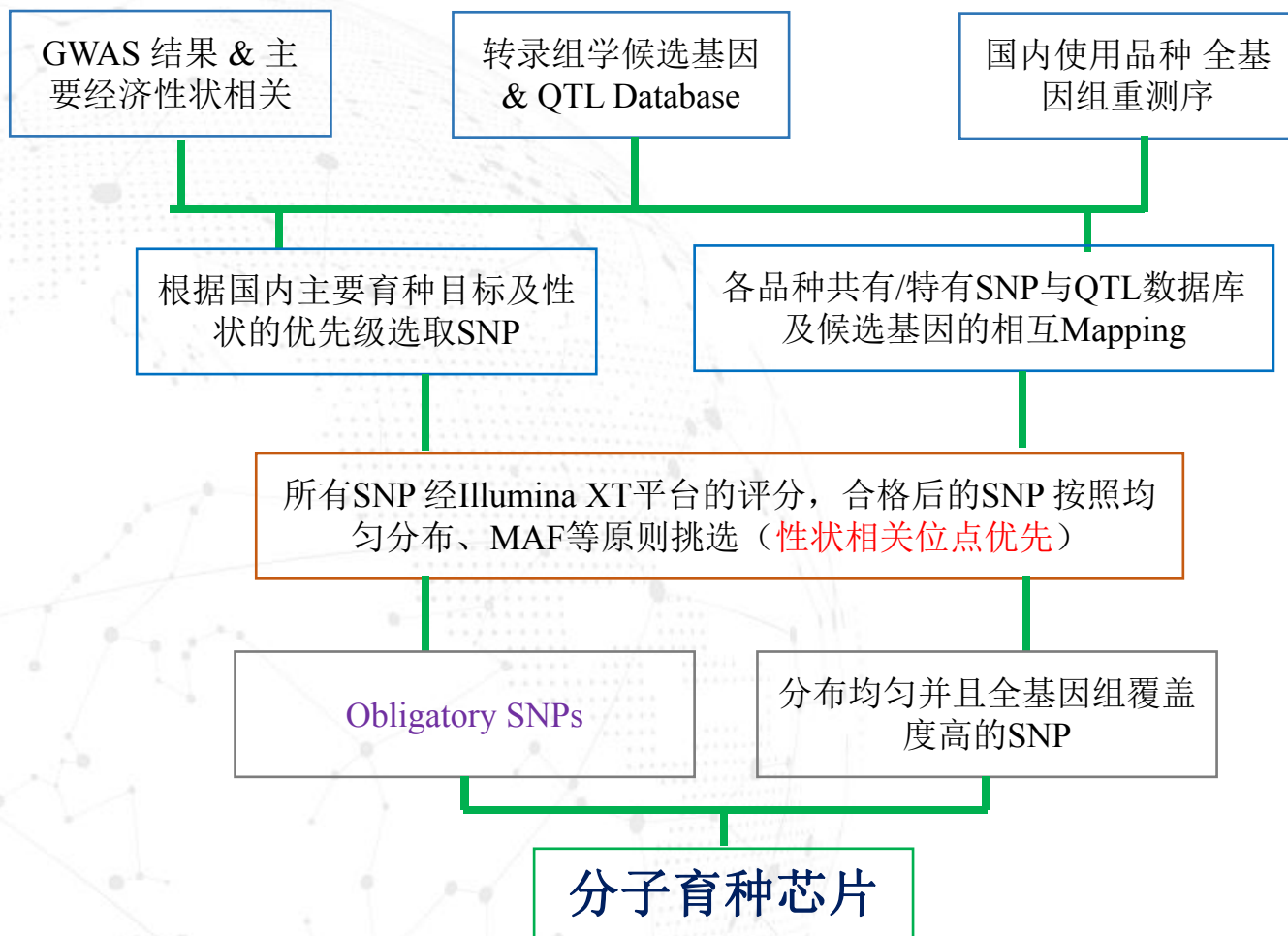
2 构建参考群及维护、算法及模型的优化

- 参考群必须和候选群体有较好的遗传联系
- 保证现有群体的全覆盖
- 表型数据准确且有高低分布，系谱要准确
- 可以进行一些模型的定制及算法的优化

3 和现有的育种体系对接

- 要符合现有企业的生产流程
- 要让一线的育种人员方便操作
- 要能及时进行相应的育种规划调整
- 场内的自己的遗传参数
- 要有自己可行的选育目标
- 要有相应的数据积累及测定设备

分子育种芯片设计



◆ 经济性状的功能相关性

生长性状、繁殖性状及健康性状等相关因果及连锁位点，增加了芯片的实用性

◆ 位点分布均匀，多态性好

筛选位点时使用重测序数据保证了位点在全基因组分布均匀且具有较高的多态性，保证了育种值估计的准确性

◆ 特异性

对比筛选了我国代表性的地方品种、国内外高度选育商用品种的特征位点，可以进行相关品种鉴定工作

◆ 先进性

整合了多个合作院所多组学和特殊资源群体的的研究成果，针对基因组上重组情况对区域位点进行了优化

GS分析主要方法及其优点

主要对比的是**BLUP**、**GBLUP**、**SSGBLUP**；养猪企业现在主要使用**SSGBLUP**方法，进行育种值估计
三个模型的主要形式：

$$y = Xb + Zu + e$$

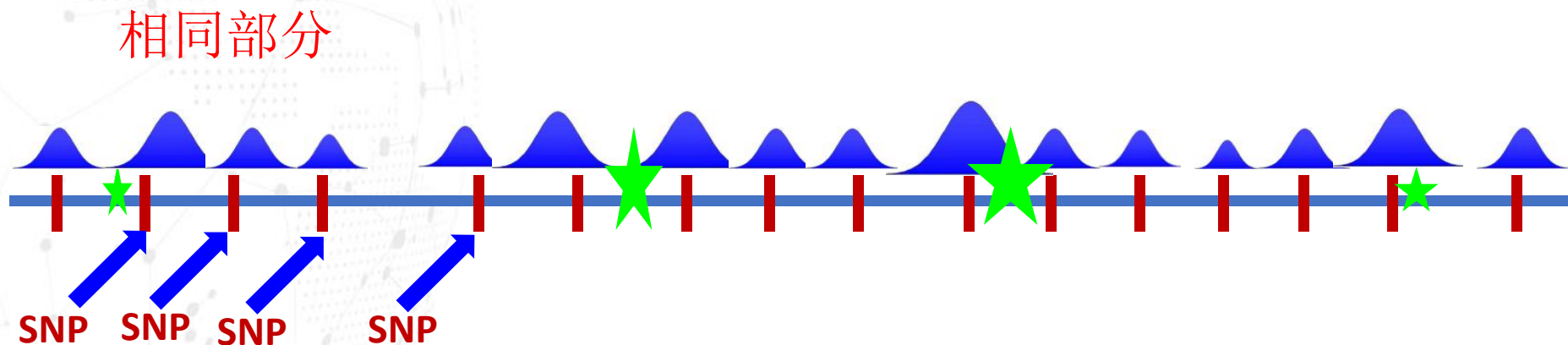
都是基于**微效多基因**理论，都是**混合线性模型**的形式
y是表型性状观察值，
b是固定效应向量，
e是随机残差效应向量， $e \sim N(0, I\sigma^2_e)$

不同的部分：**u**是随机加性遗传效应向量

BLUP:该模型中随机效应 $u \sim N(0, A\sigma^2_u)$ ，其中 **A** 为由系谱构建的分子血缘关系矩阵；

GBLUP:该模型中随机效应 $u \sim N(0, G\sigma^2_u)$ ，其中 **G** 为全基因组标记构建的基因组关系矩阵；

SSGBLUP:该模型中随机效应 $u \sim N(0, H\sigma^2_u)$ ，其中 **G** 为基因组关系矩阵和分子血缘关系矩阵合并而成；**一步法BLUP**
一步法BLUP是现在实际生产中最常用的模型

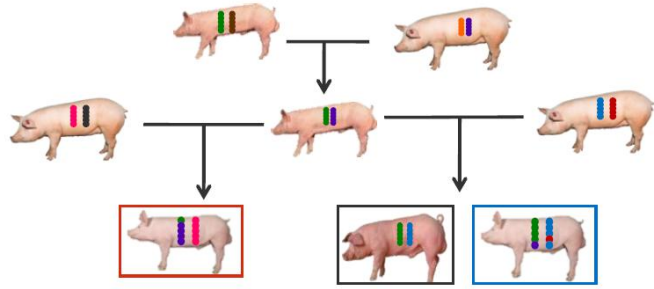


GS分析主要方法及其优点



首先BLUP VS GBLUP、SSGBLUP相比，基因组上的遗传关系相较于人为规定的系谱的遗传关系更加准确的反应的实际情况；

Relationship	Littermate	Same Sire	Grand-Sire
Assumed /假设	50%	25%	25%
True /实际	62%	9%	40%



A 矩阵	预期的亲缘关系矩阵
A matrix	Expected rel. matrix
G 矩阵	实际的亲缘关系矩阵
G matrix	Actual rel. matrix

再者，GBLUP VS SSGBLUP相比，准确性相当，但是对于实际生产应用来说性价比更高；芯片的成本使得实际生产过程中进行基因分型的个体只是一部分。

如果这个时候我们只利用G矩阵的信息，则无法进行全群的遗传评估；如果回过头利用A阵，准确性下降；所以有育种的研究者提出了G+A，重新构建H阵的构思，将常规育种和分子育种相结合。

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & t[(1 - w_1)\mathbf{G} + w_1\mathbf{A}_{22}]^{-1} - w_2\mathbf{A}_{22}^{-1} \end{bmatrix}$$

采用多性状动物模型进行分析



```
4.par
DATAFILE data.txt
NTRAITS 1
TRAITS 3
NEFFECTS 2
NVAR 3
MISSING -99
WEIGHT 0
EFFECTS
1 1/R,1 2/R,2
COV 1
1581628.4
COV 2
1617241.9
RCOV
1536088.1
COVFILE 1 INV add.inv.txt
COVFILE 2 INV epi.inv.txt
```

由于畜禽在进行选育的过程中往往不是针对单个性状。所以在分析遗传力时，常常会将两个或多个性状结合在一起进行分析，这样我们可以提高单个性状遗传参数的准确性。这时就需要用到多性状动物模型，但是由于多性状同时进行遗传评估会使模型不易收敛而且运行时间较长特别是在GS选育中。我们通过采用的Pi-BLUP可以很好的解决这些问题。

但是在进行GS分析之前，我们需要先对多性状及多性状之间的遗传相关进行方差组分求解。再得这些随机变量的方差后，不但可以进行GS的GEBV运算，还可以进行遗传力，多性状之间的表型相关和遗传相关等重要的遗传参数的估计。

这些参数对于我们制定育种目标及综合指数尤为重要。

进行模型的定制及优化

$$y=Xb+Zu+e$$

针对不同性状进行模型的选择，最常用的模型为动物模型。但是对于一些特殊的性状也会采用相应的模型比如蛋鸡的产蛋时间就会使用重复力模型。

而且随着畜禽的长久发展很多性状的评估，模型都是固定的。一般进行分析时，按照常规分析将常用的因子变量放入方程即可。

但是随着育种目标的不断丰富，更多的测定数据及环境、生产的信息不断纳入，模型可能会需要进行相应得到调整。

如果增加了新的变量，首先来看模型是否收敛，如果不收敛就要提出该效应。

如果收敛，我们再通过显著性检验来看加入的因子是否对表型真的有影响：

固定因子，我们采用**F检验**来判断，是否优化模型

随机因子，推荐使用**LRT(似然比检验)**进行相关的显著检验，

如果显著则优化模型，如果不显著则提出这个因子

通过上面可以更好的保证模型的无偏性

遗传参数和选择指数优化

➤ 遗传参数和综合选择指数的优化

通过基因分析，能够更加精确的求出不同性状的遗传方差，进而求出企业群体自身的遗传参数。

综合选择指数是企业实现自身选育目标的基础，在得到其自身场内的遗传参数后。只需客户提供我们当前的猪场群体**生产性能参数**、**管理技术参数**、**市场经济参数数**即可根据性状边际效益和性状表现贴现值的算法计算最符合客户需求的选择标准。



遗传参数具有种群特异性

Check Yourself - Duroc – heritability - daily gain

遗传参数会随着种猪遗传品质的改变而改变

加拿大种猪遗传评估体系中各主选性状的遗传力

1994年以前	日龄	背膘厚度	
约克夏	0.31	0.46	
长白猪	0.32	0.59	
汗普夏	0.35	0.44	
杜洛克	0.40	0.48	
1994年以后	日龄	背膘厚度	总产仔数
所有品种	0.30	0.52	0.11

个体综合选择指数：

$$A_T = W_1A_1 + W_2A_2 + \dots + W_kA_k = \sum W_lA_l$$

其中： A_i 是性状*i*的育种值
 W_l 是性状*i*的经济加权值

性状边际效益 (V_i)
性状表现贴现值 (n_i)

GS应用流程-样本采集及管理



样本采集、保存及运输规范

2. 耳朵样本

(1) 采样前准备:

- 配置 75%及 95%的酒精。

(2) 取样方法:

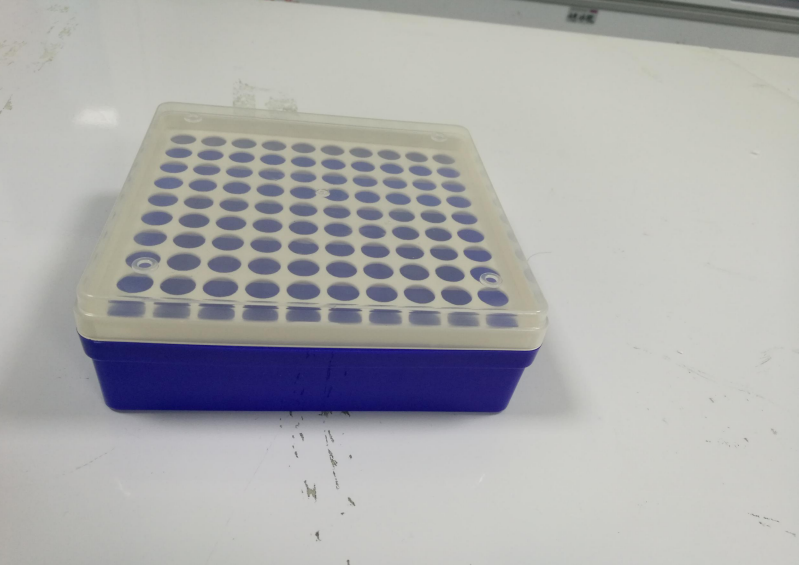
- 对耳朵用 75%酒精消毒并剪去毛发；
- 用打孔器对耳朵打孔取样（大约 100mg）；
- 将样本保存在装有 95%的酒精 EP 管中，保存于-20℃或-80℃冰箱。

(3) 样本保存及运输:

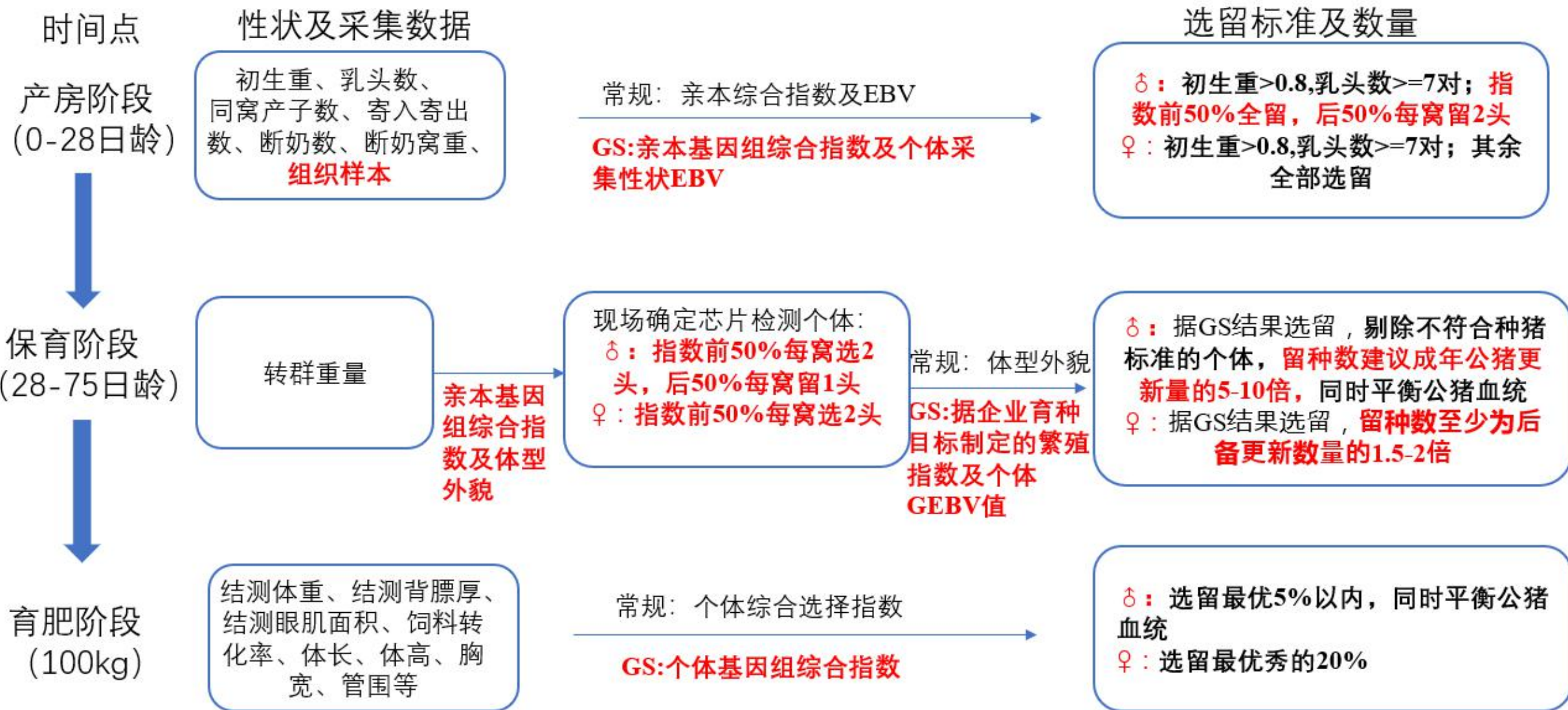
- 由于保护液为酒精，标记样本切记使用防酒精的 marker 笔；
- 干冰或冰袋运输至检测公司。



	康普森样本编号汇总表		
	文件编号:	版次: A/0	实施日期:
序号	康普森编号	原始样本名	备注
1	19023001		四川铁骑力士集团
2	19023002		四川铁骑力士集团
3	19023003		四川铁骑力士集团
4	19023004		四川铁骑力士集团
5	19023005		四川铁骑力士集团
6	19023006		四川铁骑力士集团
7	19023007		四川铁骑力士集团
8	19023008		四川铁骑力士集团
9	19023009		四川铁骑力士集团
10	19023010		四川铁骑力士集团

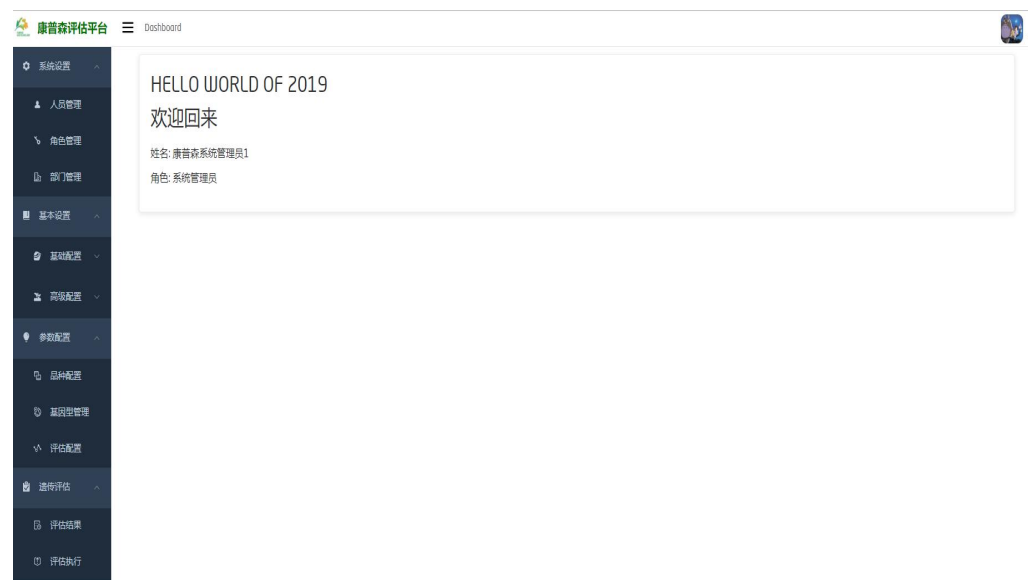


GS应用流程-检测方案制定

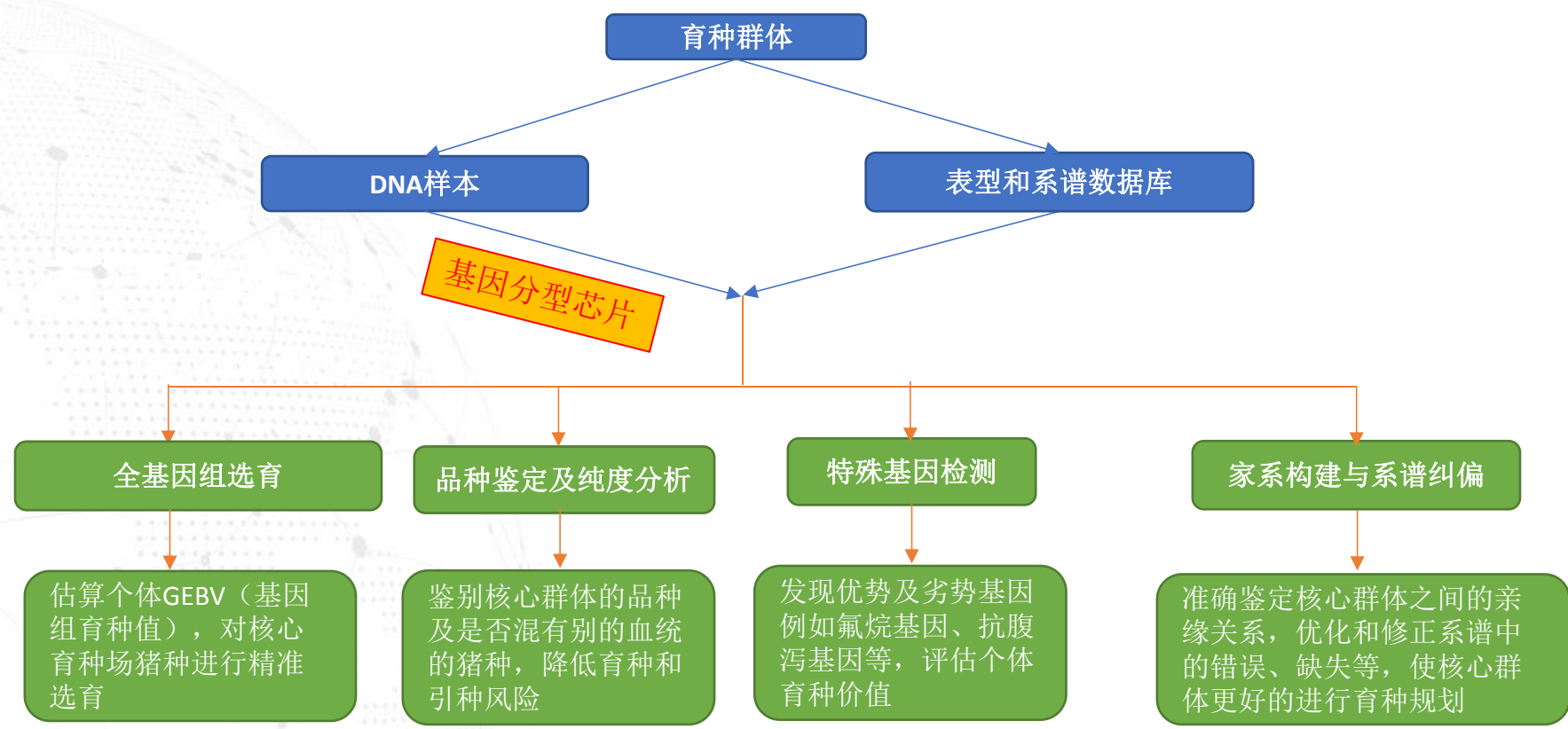


康普森遗传评估平台—和现场对接

康普森遗传评估平台是基于JavaScript和Python、MySQL等技术搭建的云端遗传评估系统，主要的运算内核基于中国农业大学开发的Pi-BLUP程序。主要目的是为企业提供高效率的育种化服务，所以其中包括企业和人员管理、育种模型制定、遗传评估（BLUP、GBLUP、SSBLUP等）、项目管理、客户上次数据逻辑检验等功能。可以有效的提升，GS项目过程中的沟通及效率。



基于芯片的育种服务

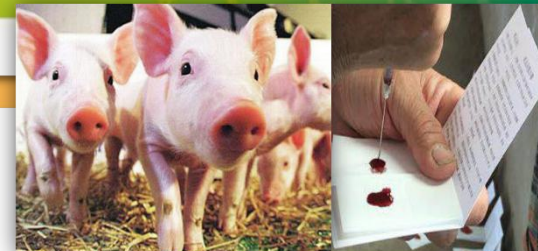


芯片进行育种服务流程



客户样本寄送与入库备份

客户将要检测的样本及样本相关数据发送北京康普森，会有专门对接的技术支持和专业的实验人员进行



DNA提取和质检

对客户提供的样品进行DNA提取及质控，保证芯片实验的成功和数据的质量



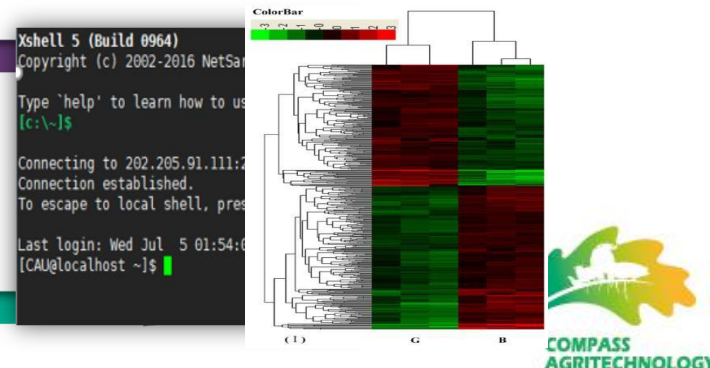
基因芯片检测实验

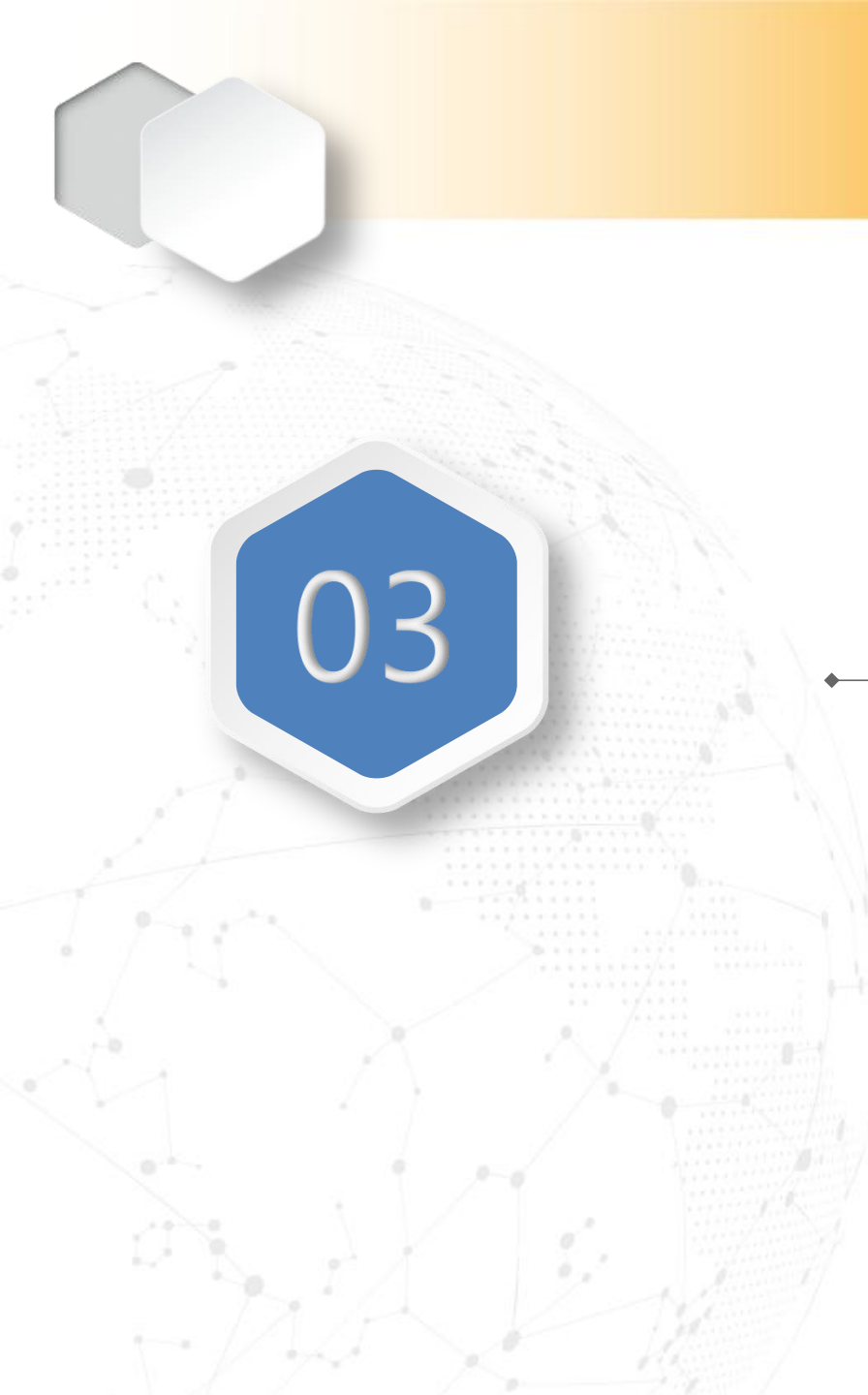
使用国际先进基因芯片检测平台，高效高质的对客户的样本进行基因分型实验



数据分析，进行分子育种

利用康普森生物数据分析集群对客户的数据进行个性化的育种分析；并按时发送结果，保证客户的育种周期



A decorative graphic on the left side of the slide. It features a large, faint network pattern of dots and lines. Overlaid on this is a blue hexagon with the number '03' in white. Above the blue hexagon are two smaller, overlapping hexagons, one grey and one white.

03

利用芯片如何进行分子育种服务

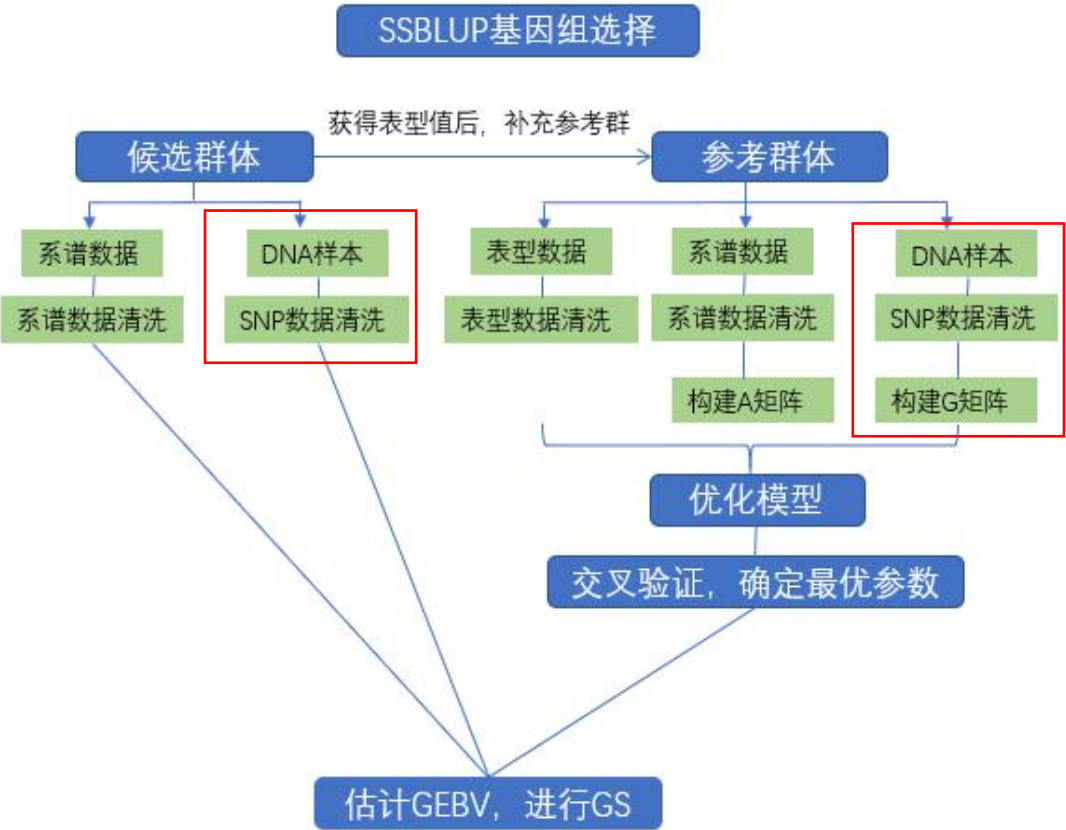
A horizontal line with diamond-shaped endpoints, separating the title from the text below.

GS为中国的企业打开了一扇大门
如何进行基因组选择？
客户能得到哪些价值？

GS分析的主要流程

一般采用SSBLUP及多性状选育的模型，芯片作为主要的基因获取工具

利用参考群的数据，进行企业自身的选育模型及H
阵构建的参数调整。



SNP data

```
1000111220020012111011112111101111001121100020122002220111
1202101200211122110021112001111001011011010220011002201101
1200201101020222121122102010011100011220221222112021120120
2010020220200002110001120201122111211102201111000021220200
0221012020002211220111012100111211102112110020102100022000
2201000201100002202211022112101121110122220012112122200200
020020202012221100222222002212111121002111120011011101120
0202220001112011010211121211102022100211201211001111102111
2110211122000101101110202200221110102011121111011202102102
1211011022122001211011211012022011002220021002110001110021
102110111000222002022121211000222010200222212121121112002
011020200122222112212021211210110012110110200220002001002
0001111011001211021212111201010121202210101011111021102112
21111121211121011012001111102111101111220121012121101022
202021211222120222002121210121210201100111222121101
```

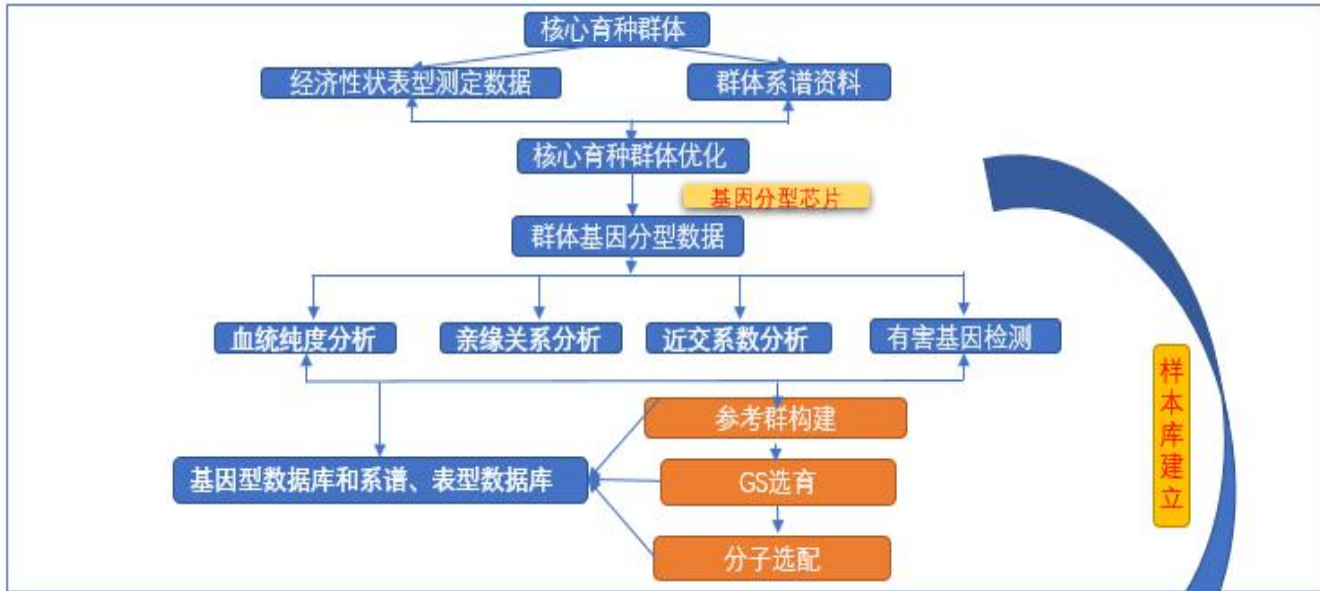
0 = homozygous for first allele (alphabetically), 1 = heterozygous
2 = homozygous for second allele (alphabetically)

$$A = b1*(0,1,2)+b2*(0,1,2)+b3*(0,1,2)+...+bn*(0,1,2)$$

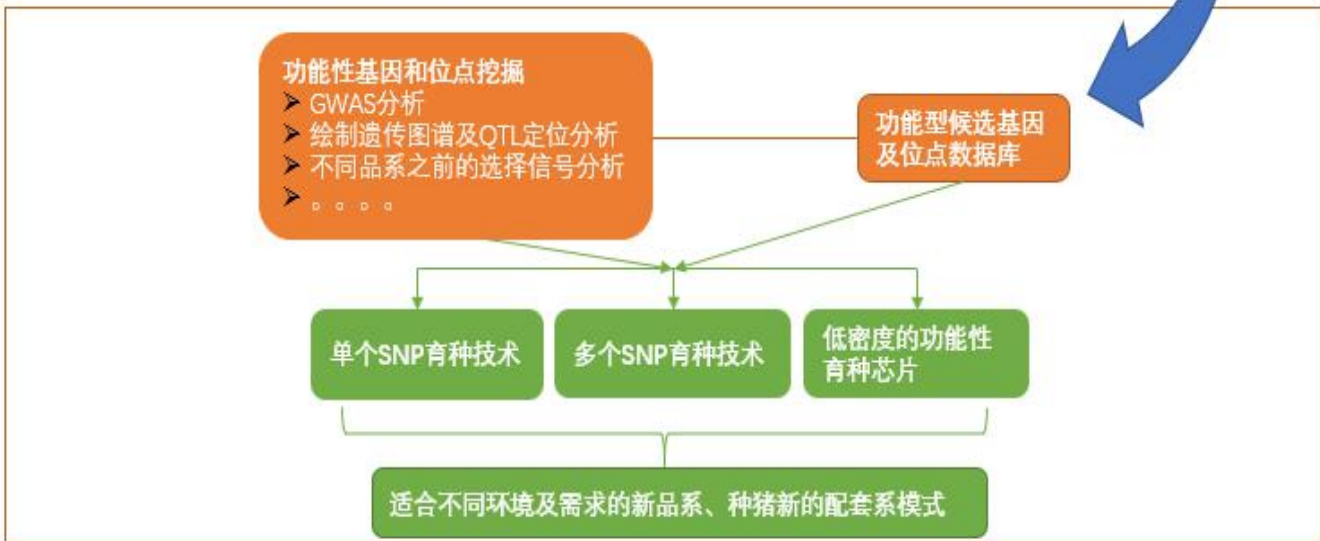
整体技术路线



第一部分



第二部分



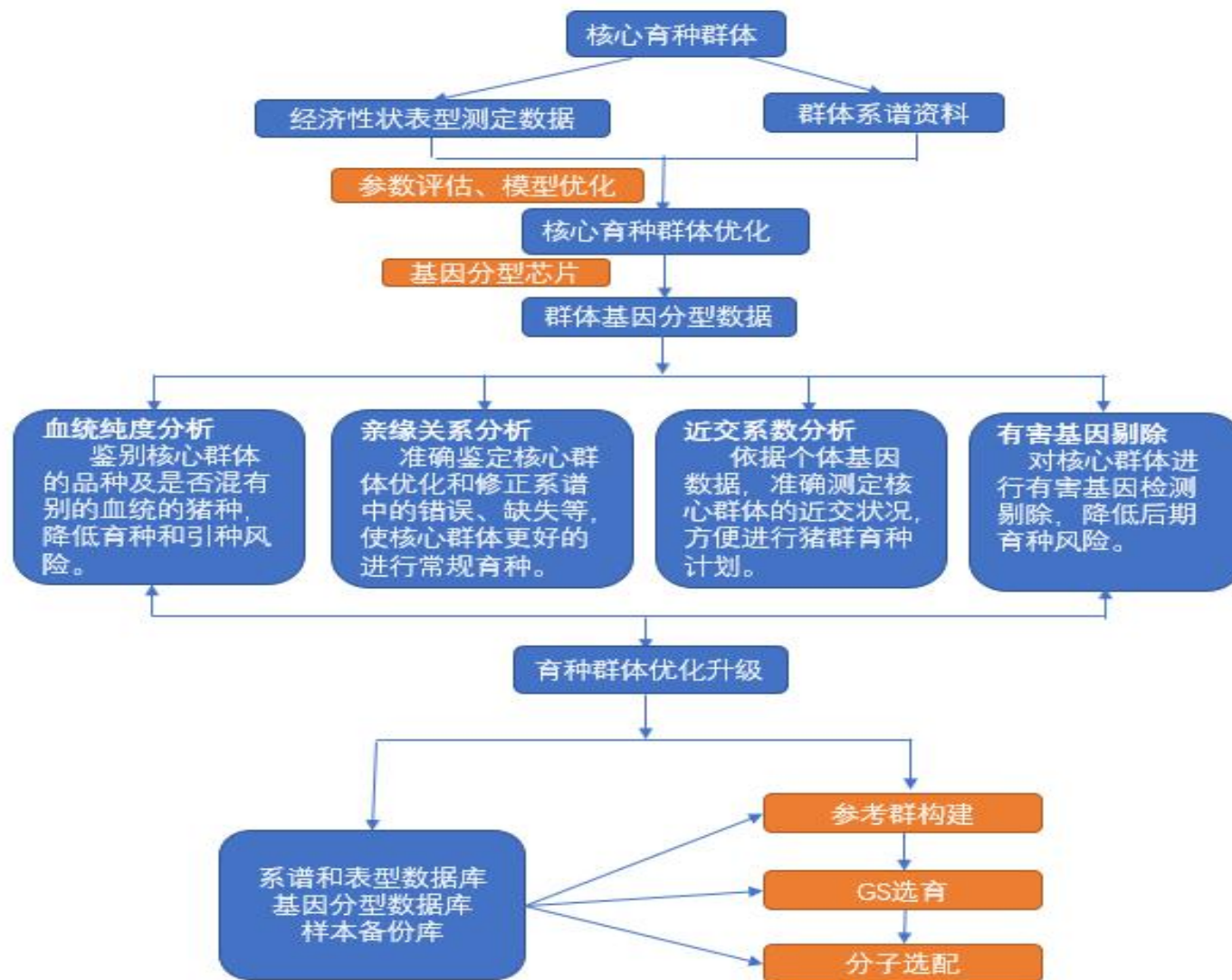
项目整体思路：应用目前前沿的技术手段结合企业育种的需求，我们以构建常规遗传评估和基因组选择全面整合的“**种猪高效分子育种体系**”为目标，进行育种板块的技术升级及研发合作。



第一阶段：针对核心育种群体，个性化定制分子育种策略



- 基于常规选育方法遗传评估优化
- 从分子层面上优化核心育种群体
- GS育种体系的建立，实现常规与分子选育的结合。
- 完成系谱、表型、样本库（功能样本库）的建立



GS分析的准备文件

	A	B	C	D	E
1	样本编号	个体管理号	品种	送样日期	备注
2	sampleID	animalID	breeds	data	note
3	1	YYTEST118000201	YY	2018/7/10	
4	2	YYTEST118001213	YY	2018/7/10	
5	3	YYTEST118000110	YY	2018/7/10	
6	4	YYTEST1180125313	YY	2018/7/10	
7	5	YYTEST118013201	YY	2018/7/10	
8	6	YYTEST118010203	YY	2018/7/10	
9	7	YYTEST118010608	YY	2018/7/10	
10					
11					
12					

样本对照表

个体管理号	父本管理号	母本管理号
animalID	sireID	damID
YYTEST118000201	YYTEST117001213	YYTEST116000110
YYTEST117001213	YYTEST1160125313	YYTEST116013201
YYTEST116000110	YYTEST115010203	YYTEST115010608
YYTEST1160125313	YYTEST215085003	YYTEST115023601
YYTEST116013201	YYTEST116000106	YYTEST115024518
YYTEST115010203	YYTEST314010215	YYTEST113018512
YYTEST115010608	YYTEST114000213	

系谱文件

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	母猪管理号	母猪品种	配种日期	公猪管理号	公猪品种	分娩日期	胎次	活仔数	死胎数	木乃伊数	出生窝重	断奶日期	断奶数	断奶窝重	场年季	其他影响因子
2	sowID	sowbreed	breeddate	boarID	boar breed	farrowdate	parity	alive num	deadbirth nu	mummy num	litter weight	weandate	wean num	wean weight	CNJ	factor
3	YYTEST116000110	YY	2018/3/10	YYTEST117001213	YY	2018/7/2	4	10	1	0	18.5	2018/7/24	10	62.1	1201803	
4	YYTEST116000111	YY	2018/3/10	YYTEST117001213	YY	2018/7/2	4	12	1	0	18.5	2018/7/24	10	62.1	1201803	
5	YYTEST116000112	YY	2018/3/10	YYTEST117001213	YY	2018/7/2	4	12	0	0	18.5	2018/7/24	10	62.1	1201803	
6	YYTEST116000113	YY	2018/3/10	YYTEST116001211	YY	2018/7/2	4	9	3	1	18.5	2018/7/24	10	62.1	1201803	
7	YYTEST116000114	YY	2018/3/10	YYTEST116001211	YY	2018/7/2	4	12	2	0	18.5	2018/7/24	10	62.1	1201803	
8	YYTEST116000115	YY	2018/3/10	YYTEST117001213	YY	2018/7/2	4	11	1	2	18.5	2018/7/24	10	62.1	1201803	
9	YYTEST116000116	YY	2018/3/10	YYTEST117001213	YY	2018/7/2	4	12	1	1	18.5	2018/7/24	10	62.1	1201803	
10	YYTEST115001710	YY	2018/3/17	YYTEST116001211	YY	2018/7/9	1	9	0	0	12.1	2018/8/1	9	62.5	1201803	
11	YYTEST115001710	YY	2018/3/18	YYTEST116001211	YY	2018/7/9	1	9	0	0	12.1	2018/8/1	9	62.5	1201803	
12	YYTEST115001710	YY	2018/3/19	YYTEST116001211	YY	2018/7/9	1	9	1	1	12.1	2018/8/1	9	62.5	1201803	

表型文件

需要提前定义好缺失值的问题（为空或者用NA）

GS应用流程-芯片检测分析

北京康普森公司 GS 分析报告

合同号: KPSNY032018023

报告单位: 北京康普森农业科技有限公司
报告日期: 2019 年 1 月 4 日
联系邮箱: agritechnology@kangpusen.com
联系电话: 010-53935082



北京康
BEIJING CC

总产仔数和产活仔数的方差组分表:

	总产仔数	产活仔数
永久性环境效应方差	1.4202134	1.0707183
加性效应方差	1.0624143	0.68909999
残差方差	10.979776	13.868041

经分析总产仔数和产活仔数的永久性环境效应方差组分分别为 1.42 和 1.07; 两个性状加性效应方差组分分别为 1.06 和 0.69; 两个性状残差方差组分分别为 10.98 和 13.87。

3.2 各性状遗传力及遗传相关

各性状遗产力

性状	遗传力	遗传相关
百公斤日龄	0.324089	-0.19327
百公斤背膘厚	0.5319	
总产仔数	0.078917	
产活仔数	0.044094	

结论: 在进行方差组分估计是通过利用参考群数据进行计算, 用于方差组分估计的生长数据性状共 26579 条, 繁殖数据 6727 条, 通过结果可以看出各个性状遗传力均在正常范围内, 百公斤日龄为中等遗传力、百公斤背膘厚为中高遗传力而产仔性状为低遗传力。

GS优势-早期选择，降低成本



选育技术	初选 (10日龄)	第二次选择 (21-45days)	终测数量 (150days)	留种数量	技术优势
PBLUP	留50%，约 3600	1500 (♂)	1200 (♂)	35	150日龄测定后方可进行种猪选留
	7000	2800 (♀)	2400 (♀)	400	
SSBLUP	留50%，约 3600	1800 (♂)	300 (♂)	35	在45日龄前即可进行种猪选留工作
	7000	1200 (♀)	800 (♀)	400	

GS优势-早期选择，降低成本



GS技术降低测定成本：

公猪淘汰收入方面：在国内淘汰公猪的价格一般在肥猪价格的三分之二，如按肥猪价格14元/kg,则每头肥猪损失约500元，因此通过早期选择可增加公猪收入 $(1500-300) \times 500 = 60$ 万。

饲料成本方面：通过GS早期选择后，不进行留种的种猪可在60kg销售，比常规100公斤测定后销售约节省饲料100 公斤，每公斤饲料2.5元计算。即 $(3000-300-800) \times 100 \times 2.5 = 47.5$ 万元；

猪只疫苗成本降低：通过早期种猪选择可减少疫苗成本约 $(3000-300-800) \times 25 = 4.75$ 万元

猪只损失降低：

人工饲养成本降低：

栋舍利用率提高：

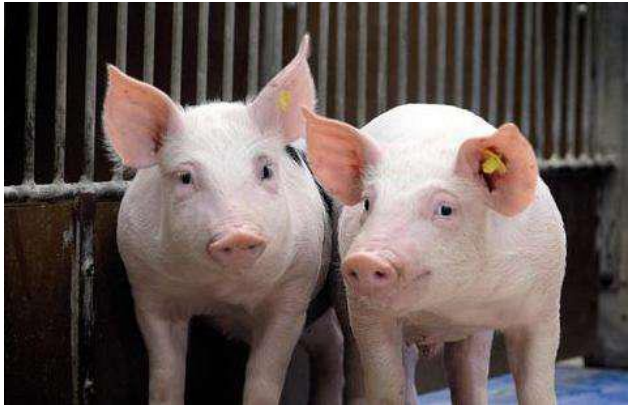
GS技术投入： $3000 \times 198 = 59.4$ 万元

综上所述：育种带来的遗传进展是可累加的，因此GS技术给企业带来的效益远远大于投入。

基于芯片的特殊基因（因果位点）检测

基因	QQ	Qq	qq	效应	育种建议
肌间脂肪	260	0	0	Q等位基因有利于肌间脂肪增加	已选纯
脂肪酸组成	260	0	0	Q等位基因增加C18:1比例,降低C18：0比例	已选纯
酸肉基因	260	0	0	q等位基因导致酸肉	已选纯
氟烷基基因	235	25	0	q等位基因造成应激	有25头杂合子,可以根据血缘关系迅速选纯
多肋骨数基因	0	84	176	Q等位基因型增加0.5根胸椎	无QQ纯合个体，需要从84个杂合子中产生QQ纯合子，考虑血缘情况下增加Q的频率
体长基因	0	24	236	Q等位基因型增加5.8厘米体长	无QQ纯合个体，需要从24个杂合子中产生QQ纯合子，考虑血缘情况下增加Q的频率
抗仔猪腹泻基因	130	90	40	QQ基因型具抗仔猪腹泻潜力	尽量提升QQ个体频率

让客户更加了解自身群体的种质资源。通过育种技术降低有害基因及提升有利基因频率。后期特殊基因的群体还可以用于专业商品化品系及配套系的研发。特别在当前非洲猪瘟较为严重的时期，可以让企业通过冻精、生物克隆、生产管理等方面提前对优势基因的个体进行保护。



基于芯片技术的系谱纠偏



- ✓ 引种时系谱混乱
- ✓ 现场种畜耳标脱落，身份不明
- ✓ 人工授精操作不规范
- ✓ 人为记录的失误

一般在进行SSBLUP过程中，会对有基因型的个体进行系谱检验。这些有基因型的个体后期大多会留种，所以开展GS的企业能大大降低系谱错误对企业生产造成的影响。

孟德尔错误原理

基于孟德尔遗传定律，即每个遗传位点的等位基因均以孟德尔遗传方式由亲本传递给后代。在疑似亲本和疑似后代个体所构成的待检测亲子对间，可对每一个双等位基因的遗传位点用如下规则进行孟德尔错误判定：

理论上只要100个左右合适的标记，就能得到准确率为99.99%以上的判断。

表 4. 亲子鉴定结果

Offspring ID	Candidate father ID	Candidate mother ID
1		CNXD001201630674
3	CNXD001201630052	CNXD001201630643
4	CNXD001201630293	CNXD001201630449
5	CNXD001201630214	CNXD001201630015
6	CNXD001201630277	CNXD001201630674
8	CNXD001201630265	CNXD001201630449
9		CNXD001201630643
10	FR56TCE201502529	CNXD001201630015
11	CNXD001201630052	CNXD001201630643

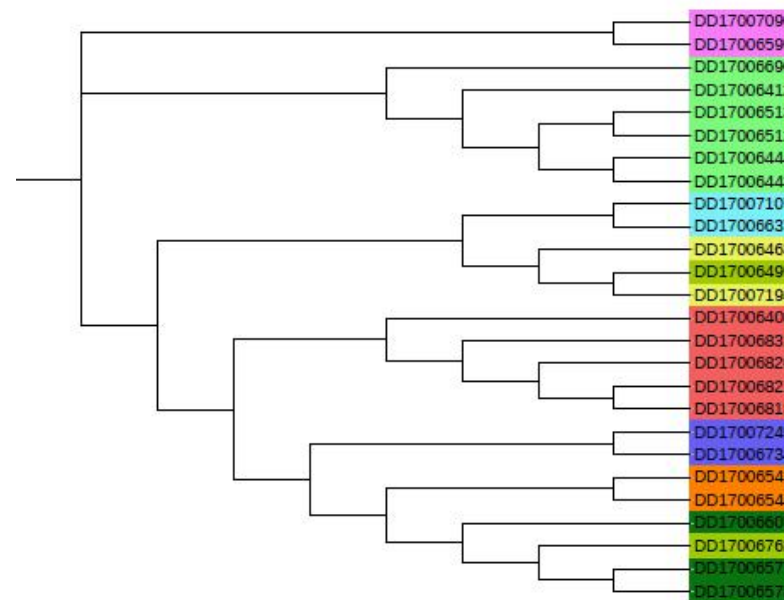
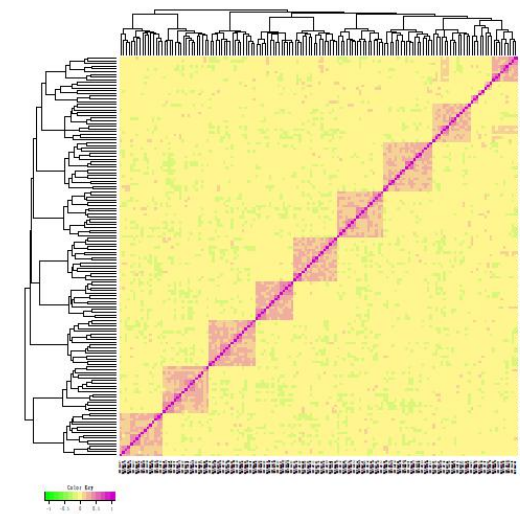
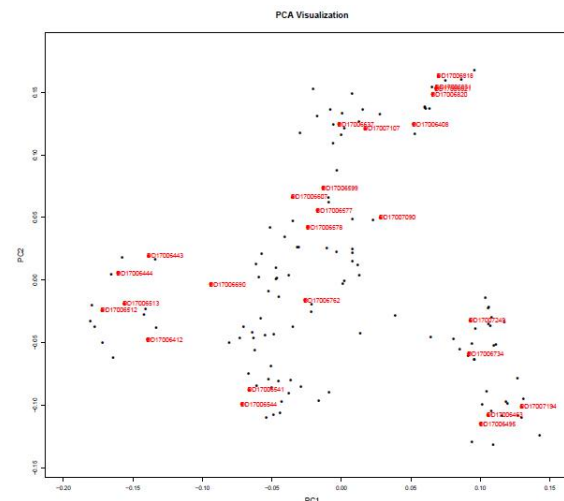
基于芯片的家系构建

- ✓ 解析群体中个体间的基因组亲缘关系
- ✓ 为后续配种方案提供参考依据
- ✓ 帮助企业科学的构建家系
- ✓ 所有分析均基于全基因组数据，降低系谱不全或错误带来的分析误差

$$G = \frac{ZZ'}{2 \sum p_i(1 - p_i)}$$

$$G_{ij} = \sqrt{G_{ii} * G_{jj}}$$

VanRaden 2008 Efficient Methods to Compute Genomic Predictions



基于基因组亲缘关系的家系构建



客户引入种猪时，有12个血统。而从基因组亲缘关系上，该群体只有8个左右的家系。

这样可以让客户在制定育种计划时，更加科学的减少近交增量。

对于实行GS的企业来说，该技术还能应用于纯系的选配环节，进一步的降低近交增量的问题。

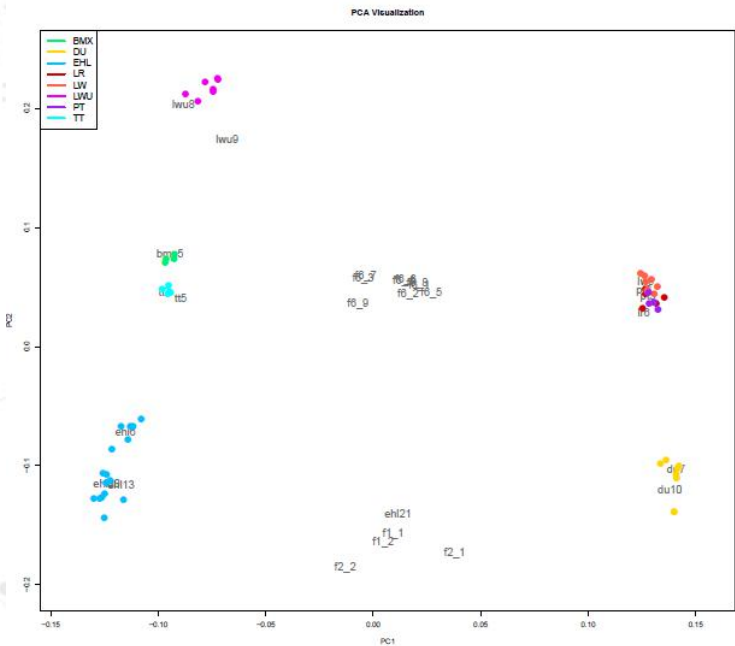
家系A	公猪（2头）	DD17007090 DD17006599					
	母猪（14头）	DD17007094	DD17007093	DD17006979	DD17006982	DD17006977	DD17006602
		DD17006605	DD17007507	DD17006478	DD17007182	DD17006421	DD17007183
		DD17007506	DD17006813				
家系B	公猪（6头）	DD17006513	DD17006444	DD17006443	DD17006690	DD17006512	DD17006412
	母猪（14头）	DD17006517	DD17007573	DD17006922	DD17006750	DD17007529	DD17006447
		DD17006448	DD17006748	DD17006751	DD17006996	DD17006747	DD17006668
		DD17006482	DD17007568				
家系C	公猪（2头）	DD17006637	DD17007107				
	母猪（16头）	DD17006639	DD17006640	DD17006572	DD17006594	DD17007229	DD17007133
		DD17007111	DD17006668	DD17007231	DD17007132	DD17007134	DD17006652
		DD17007182	DD17007183	DD17006482	DD17006478		
家系D	公猪（3头）	DD17006463	DD17006495	DD17007194			
	母猪（14头）	DD17006467	DD17006466	DD17006620	DD17007103	DD17007064	DD17007200
		DD17007298	DD17006500	DD17007104	DD17007203	DD17006482	DD17006499
		DD17007302	DD17007199				
家系E	公猪（5头）	DD17006820	DD17006821	DD17006831	DD17006818	DD17006408	
	母猪（10头）	DD17006822	DD17006825	DD17006834	DD17007520	DD17007521	DD17006833
		DD17006405	DD17006832	DD17006813	DD17006594		
家系F	公猪（2头）	DD17006734	DD17007249				
	母猪（17头）	DD17006915	DD17006908	DD17007236	DD17007088	DD17006539	DD17007086
		DD17006459	DD17006540	DD17006904	DD17007252	DD17006439	DD17006457
		DD17006706	DD17006708	DD17006458	DD17006438	DD17006801	
家系G	公猪（2头）	DD17006544	DD17006541				
	母猪（16头）	DD17006547	DD17006549	DD17007125	DD17007268	DD17007165	DD17007364
		DD17006555	DD17007266	DD17006554	DD17007162	DD17007163	DD17006559
		DD17006556	DD17007158	DD17007366	DD17006807		
家系H	公猪（4头）	DD17006577	DD17006762	DD17006578	DD17006607		
	母猪（13头）	DD17006584	DD17006678	DD17006877	DD17006801	DD17006798	DD17006614
		DD17006807	DD17006900	DD17006618	DD17006892	DD17006616	DD17006802
		DD17006594					
其他	母猪（17头）	DD17007035	DD17007411	DD17007607	DD17007451	DD17007207	DD17007434



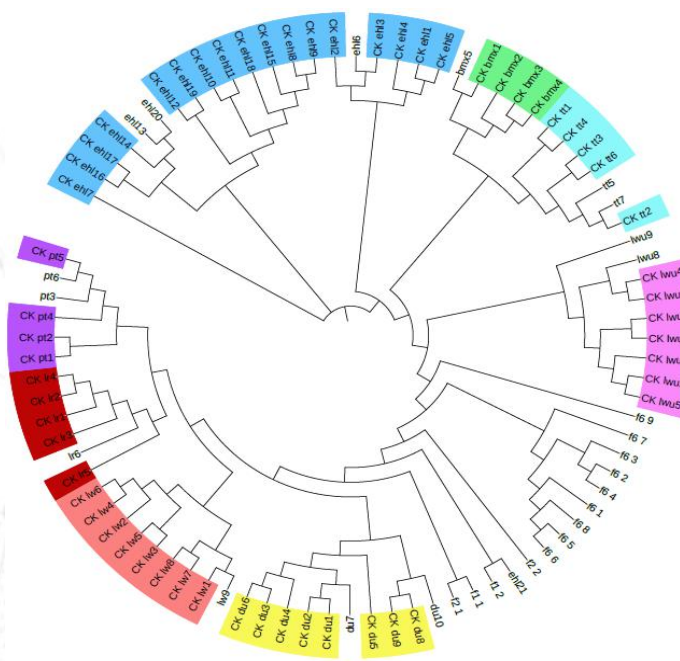
基于芯片血统纯度分析



PCA



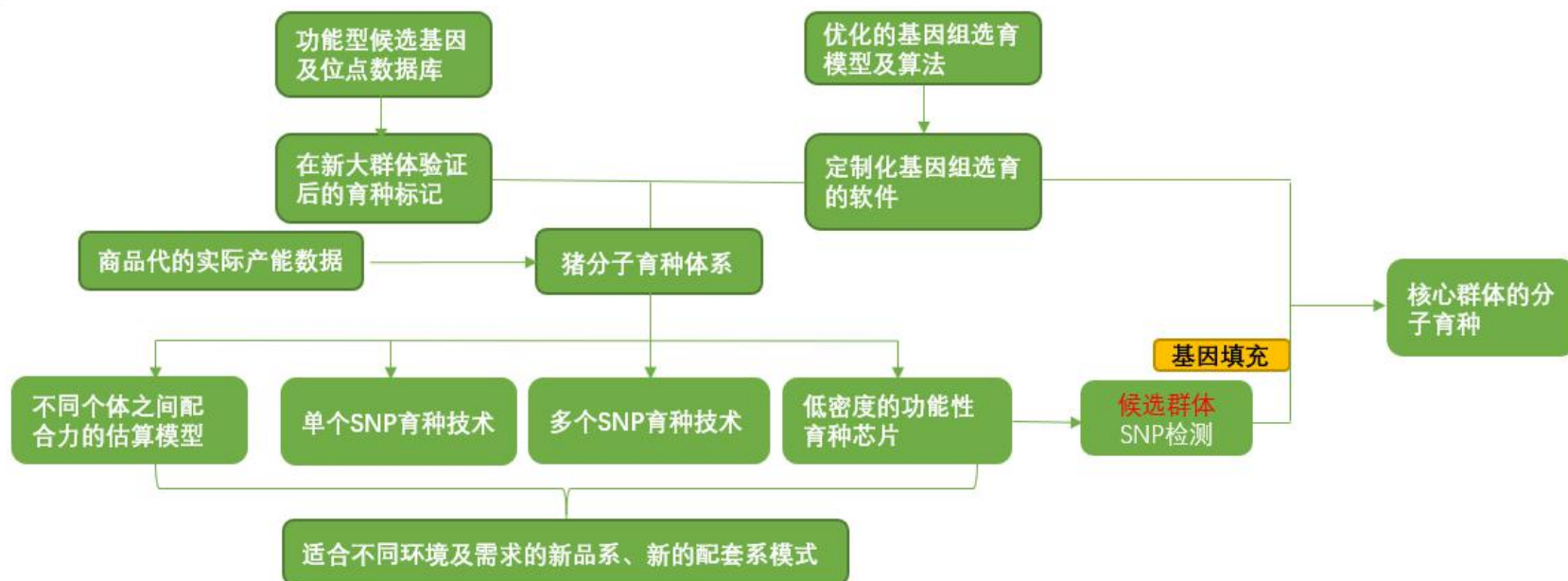
聚类分析



全基因组遗传相似度分析

	bmx5	du7	eh16	eh121
bmx5	1	0.3617	0.5963	0.4157
du7	0.3617	1	0.3479	0.4468
eh16	0.5963	0.3479	1	0.458
eh121	0.4157	0.4468	0.458	1
CK_bmx1	0.8442	0.3635	0.595	0.4178
CK_bmx2	0.7194	0.3533	0.6053	0.4144
CK_bmx3	0.7325	0.3606	0.5968	0.4217
CK_bmx4	0.7158	0.3549	0.5993	0.4209
CK_du1	0.3592	0.7572	0.3483	0.4442
CK_du2	0.3569	0.7506	0.3474	0.444
CK_du3	0.356	0.6533	0.351	0.4518
CK_du4	0.3571	0.6292	0.3533	0.4527
CK_eh12	0.5942	0.3565	0.7232	0.4631
CK_eh13	0.5943	0.3451	0.7512	0.4553
CK_eh14	0.5983	0.3549	0.6989	0.4659

第二阶段：针对企业产业布局，搭建“高效分子育种体系”



- 利用生物信息技术对重要性状的关联基因、位点进行挖掘，构建企业自身群体的功能性基因及标记数据库
- 合作开发，完成“高效的分子育种体系”搭建
- 建立自主知识产权育种分子标记库，引领国内育种方向，提升行业影响力。
- 打造标记辅助选择、专业化商品猪分子选育、育种软件、分子选配等高效分子育种体系，加速自主种猪新品系和新配套系开发。



THANK YOU

北京康普森农业科技有限公司
BEIJING COMPASS AGRITECHNOLOGY CO., LTD

010-53935082

📍 北京市昌平区TBD云集中心2号楼C区608
🌐 www.kangpusen.com
✉ agritechnology@kangpusen.com

