**分子标记检测报告**

# 1. 研究背景

烟草垄叶数基因位点验证，对群体的垄叶数极端个体，利用 6个分子标记使用单因素方差分析验证其位点表型相关性。

# 2. 材料与方法

## 2.1 试验材料

本项目群体包含同时拥有基因型和垄叶数目表型的 869 个单株。

## 2.2 标记开发检测结果

对6个分子标记进行单因素方差分析检测。其中，6个标记（K000754、K000782、K000783、SNPY3-5）均显示极显著关联（p<1e-16），F值范围38.42-41.68（表明基因型分组可解释12%-15%表型变异。其中SNPY4效应最强（F=41.68）），F值是组间方差（MSB）与组内方差（MSW）的比值，F值大于1，可说明组间差异由处理效应引起。P值越小，F值越大，则能更有力地支持拒绝H₀的结论。

### 分子标记K000754

**（1） ANOVA分析**：单因素方差分析是一种用于比较三个或更多组的均值是否存在显著差异的统计方法，适用于单一分类变量（自变量）对连续型因变量的影响分析。

**表1：单因素方差分析结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | ddof1 | ddof2 | F | p-unc |
| group | 2 | 872 | 3.842e+01 | 1.017e-16 |

Source :变异来源 : "group":（组间变异）,"Residual"或"Error":(未被解释的随机变异),"Total":(整体变异)

ddof1 :组间自由度: 自由度=分组数-1 : 值为2 代表3个分组

ddof2 :组内自由度: 自由度=组内自由度+组间自由度+1 ：值为845 表示总样本量为845,反映独立信息量

F :F统计量 : 组间组内方差比 : 值为58,则表示组间方差是组内方差的58倍

p-unc :未校正p值 : 结果是偶然发生的概率(无效假设)，P值远小于常规阈值0.05 : 拒绝假设

**（2） Tukey HSD**（Honestly Significant Difference，真实显著差异）检验是用于多重比较的统计方法，专门解决ANOVA（方差分析）后多组间两两比较的假阳性问题。

**表2：多重比较检验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A | B | mean(A) | mean(B) | diff | p-tukey | sig |
| CC | TC | 2.527e+01 | 2.684e+01 | -1.568e+00 | 1.000e-03 | \*\* |
| CC | TT | 2.527e+01 | 2.792e+01 | -2.649e+00 | 1.000e-03 | \*\* |
| TC | TT | 2.684e+01 | 2.792e+01 | -1.081e+00 | 1.000e-03 | \*\* |

A : 第一个比较组名

B : 第二个比较组名,AB先后顺序无影响

mean(A) : 第一组均值

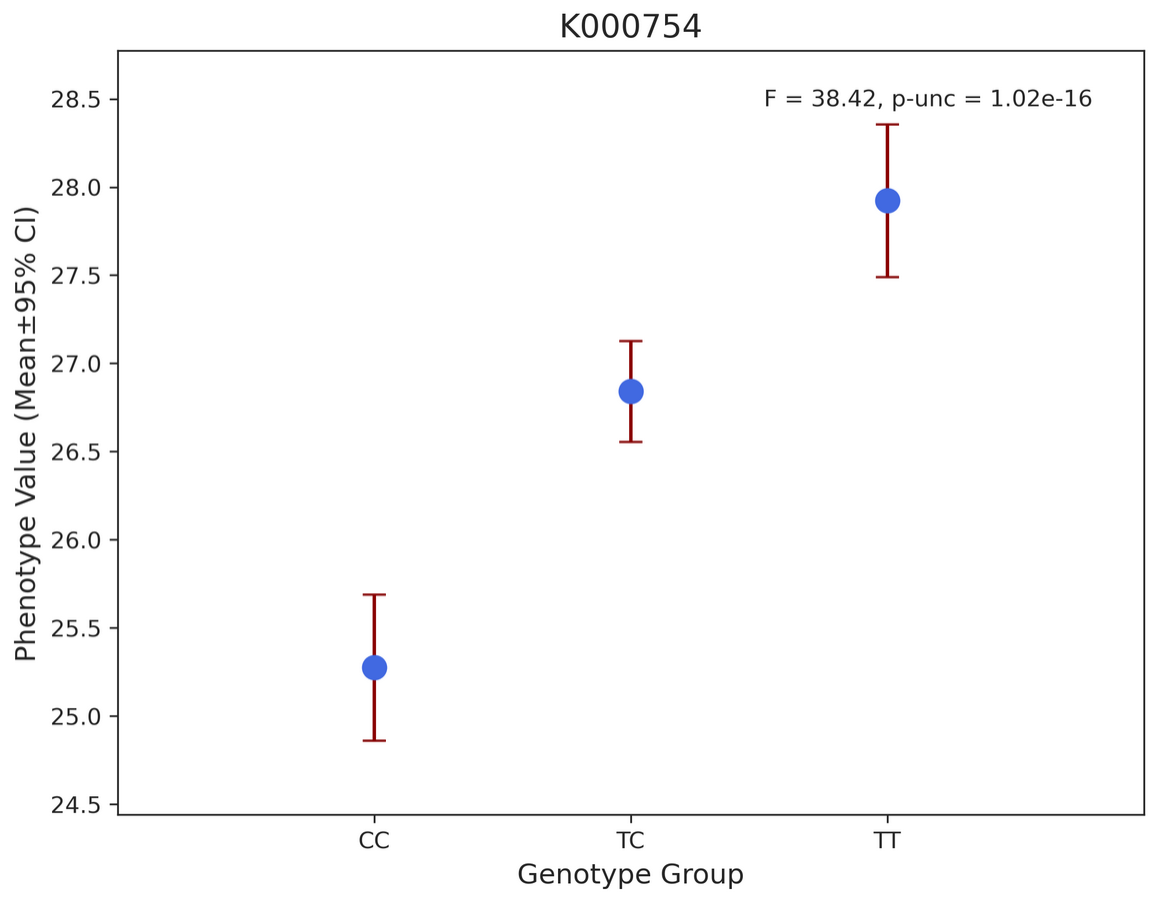
mean(B) : 第二组均值

diff : 组间均值差

p-tukey : 校正后P值,通过Tukey方法调整多重比较的显著性

sig : 显著性标记 p < 0.01（高度显著\*\*）p < 0.05（显著\*）

**（3） 组间均值比较图**



**图1：组间均值比较**

X轴：分组类别；Y轴：表型值均值；误差线：95%置信区间；图中标注F统计量和p值

### 分子标记K000782

**（1） ANOVA分析**：单因素方差分析是一种用于比较三个或更多组的均值是否存在显著差异的统计方法，适用于单一分类变量（自变量）对连续型因变量的影响分析。

**表1：单因素方差分析结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | ddof1 | ddof2 | F | p-unc |
| group | 2 | 864 | 4.076e+01 | 1.217e-17 |

Source :变异来源 : "group":（组间变异）,"Residual"或"Error":(未被解释的随机变异),"Total":(整体变异)

ddof1 :组间自由度: 自由度=分组数-1 : 值为2 代表3个分组

ddof2 :组内自由度: 自由度=组内自由度+组间自由度+1 ：值为845 表示总样本量为845,反映独立信息量

F :F统计量 : 组间组内方差比 : 值为58,则表示组间方差是组内方差的58倍

p-unc :未校正p值 : 结果是偶然发生的概率(无效假设)，P值远小于常规阈值0.05 : 拒绝假设

**（2） Tukey HSD**（Honestly Significant Difference，真实显著差异）检验是用于多重比较的统计方法，专门解决ANOVA（方差分析）后多组间两两比较的假阳性问题。

**表2：多重比较检验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A | B | mean(A) | mean(B) | diff | p-tukey | sig |
| AA | GA | 2.517e+01 | 2.689e+01 | -1.728e+00 | 1.000e-03 | \*\* |
| AA | GG | 2.517e+01 | 2.790e+01 | -2.736e+00 | 1.000e-03 | \*\* |
| GA | GG | 2.689e+01 | 2.790e+01 | -1.008e+00 | 1.000e-03 | \*\* |

A : 第一个比较组名

B : 第二个比较组名,AB先后顺序无影响

mean(A) : 第一组均值

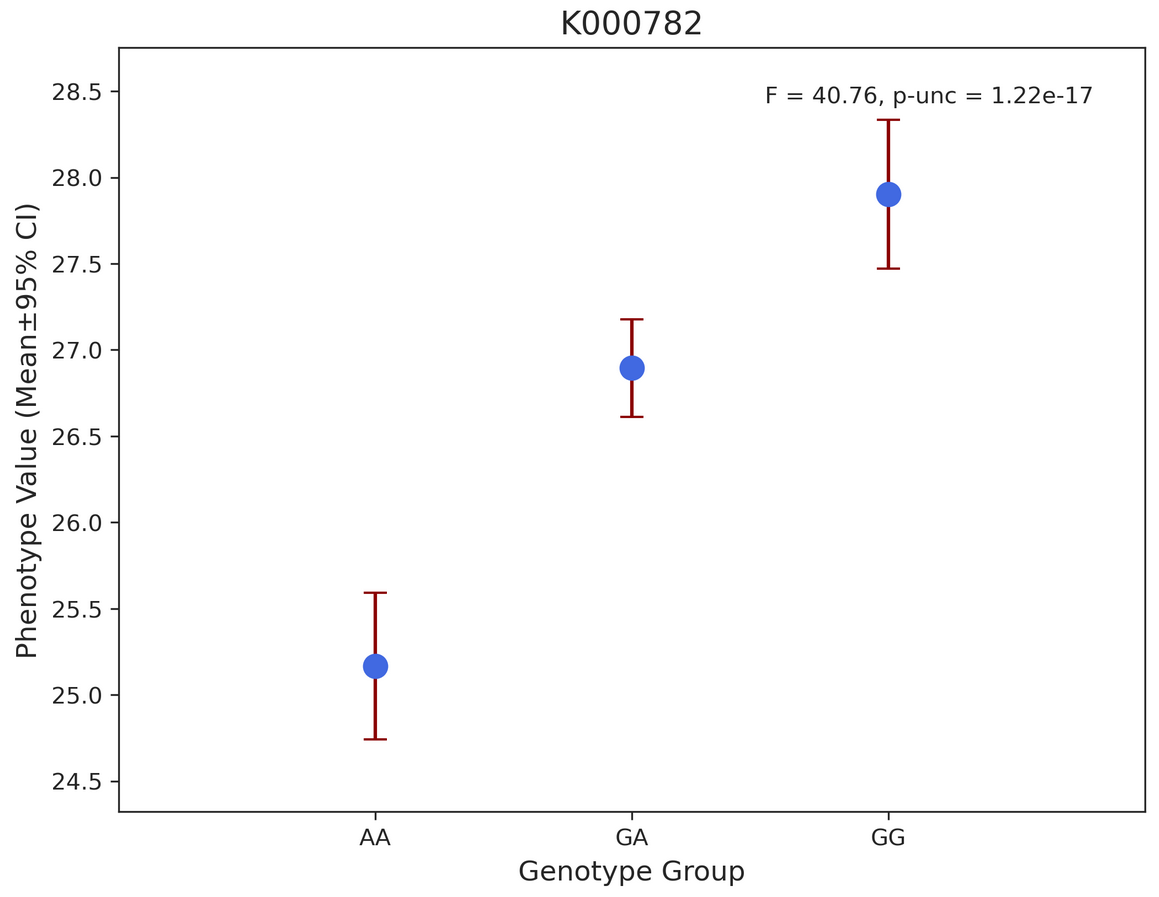
mean(B) : 第二组均值

diff : 组间均值差

p-tukey : 校正后P值,通过Tukey方法调整多重比较的显著性

sig : 显著性标记 p < 0.01（高度显著\*\*）p < 0.05（显著\*）

**（3） 组间均值比较图**



**图1：组间均值比较**

X轴：分组类别；Y轴：表型值均值；误差线：95%置信区间；图中标注F统计量和p值

### 分子标记K000783

**（1） ANOVA分析**：单因素方差分析是一种用于比较三个或更多组的均值是否存在显著差异的统计方法，适用于单一分类变量（自变量）对连续型因变量的影响分析。

**表1：单因素方差分析结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | ddof1 | ddof2 | F | p-unc |
| group | 2 | 864 | 3.998e+01 | 2.474e-17 |

Source :变异来源 : "group":（组间变异）,"Residual"或"Error":(未被解释的随机变异),"Total":(整体变异)

ddof1 :组间自由度: 自由度=分组数-1 : 值为2 代表3个分组

ddof2 :组内自由度: 自由度=组内自由度+组间自由度+1 ：值为845 表示总样本量为845,反映独立信息量

F :F统计量 : 组间组内方差比 : 值为58,则表示组间方差是组内方差的58倍

p-unc :未校正p值 : 结果是偶然发生的概率(无效假设)，P值远小于常规阈值0.05 : 拒绝假设

**（2） Tukey HSD**（Honestly Significant Difference，真实显著差异）检验是用于多重比较的统计方法，专门解决ANOVA（方差分析）后多组间两两比较的假阳性问题。

**表2：多重比较检验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A | B | mean(A) | mean(B) | diff | p-tukey | sig |
| AA | AT | 2.782e+01 | 2.693e+01 | 8.908e-01 | 1.684e-03 | \*\* |
| AA | TT | 2.782e+01 | 2.515e+01 | 2.669e+00 | 1.000e-03 | \*\* |
| AT | TT | 2.693e+01 | 2.515e+01 | 1.778e+00 | 1.000e-03 | \*\* |

A : 第一个比较组名

B : 第二个比较组名,AB先后顺序无影响

mean(A) : 第一组均值

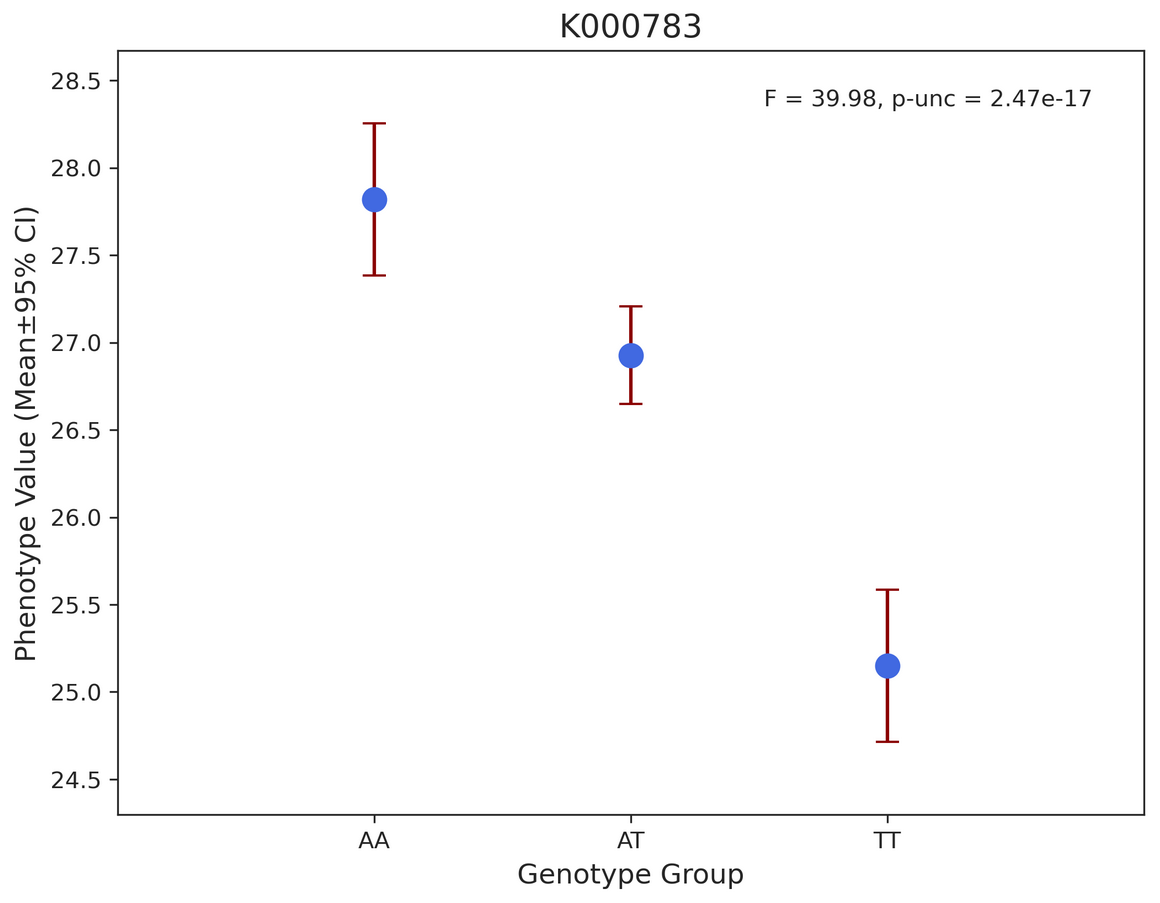
mean(B) : 第二组均值

diff : 组间均值差

p-tukey : 校正后P值,通过Tukey方法调整多重比较的显著性

sig : 显著性标记 p < 0.01（高度显著\*\*）p < 0.05（显著\*）

**（3） 组间均值比较图**



**图1：组间均值比较**

X轴：分组类别；Y轴：表型值均值；误差线：95%置信区间；图中标注F统计量和p值

### 分子标记SNPY3

**（1） ANOVA分析**：单因素方差分析是一种用于比较三个或更多组的均值是否存在显著差异的统计方法，适用于单一分类变量（自变量）对连续型因变量的影响分析。

**表1：单因素方差分析结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | ddof1 | ddof2 | F | p-unc |
| group | 2 | 872 | 4.063e+01 | 1.347e-17 |

Source :变异来源 : "group":（组间变异）,"Residual"或"Error":(未被解释的随机变异),"Total":(整体变异)

ddof1 :组间自由度: 自由度=分组数-1 : 值为2 代表3个分组

ddof2 :组内自由度: 自由度=组内自由度+组间自由度+1 ：值为845 表示总样本量为845,反映独立信息量

F :F统计量 : 组间组内方差比 : 值为58,则表示组间方差是组内方差的58倍

p-unc :未校正p值 : 结果是偶然发生的概率(无效假设)，P值远小于常规阈值0.05 : 拒绝假设

**（2） Tukey HSD**（Honestly Significant Difference，真实显著差异）检验是用于多重比较的统计方法，专门解决ANOVA（方差分析）后多组间两两比较的假阳性问题。

**表2：多重比较检验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A | B | mean(A) | mean(B) | diff | p-tukey | sig |
| CC | CT | 2.788e+01 | 2.691e+01 | 9.617e-01 | 1.000e-03 | \*\* |
| CC | TT | 2.788e+01 | 2.519e+01 | 2.685e+00 | 1.000e-03 | \*\* |
| CT | TT | 2.691e+01 | 2.519e+01 | 1.724e+00 | 1.000e-03 | \*\* |

A : 第一个比较组名

B : 第二个比较组名,AB先后顺序无影响

mean(A) : 第一组均值

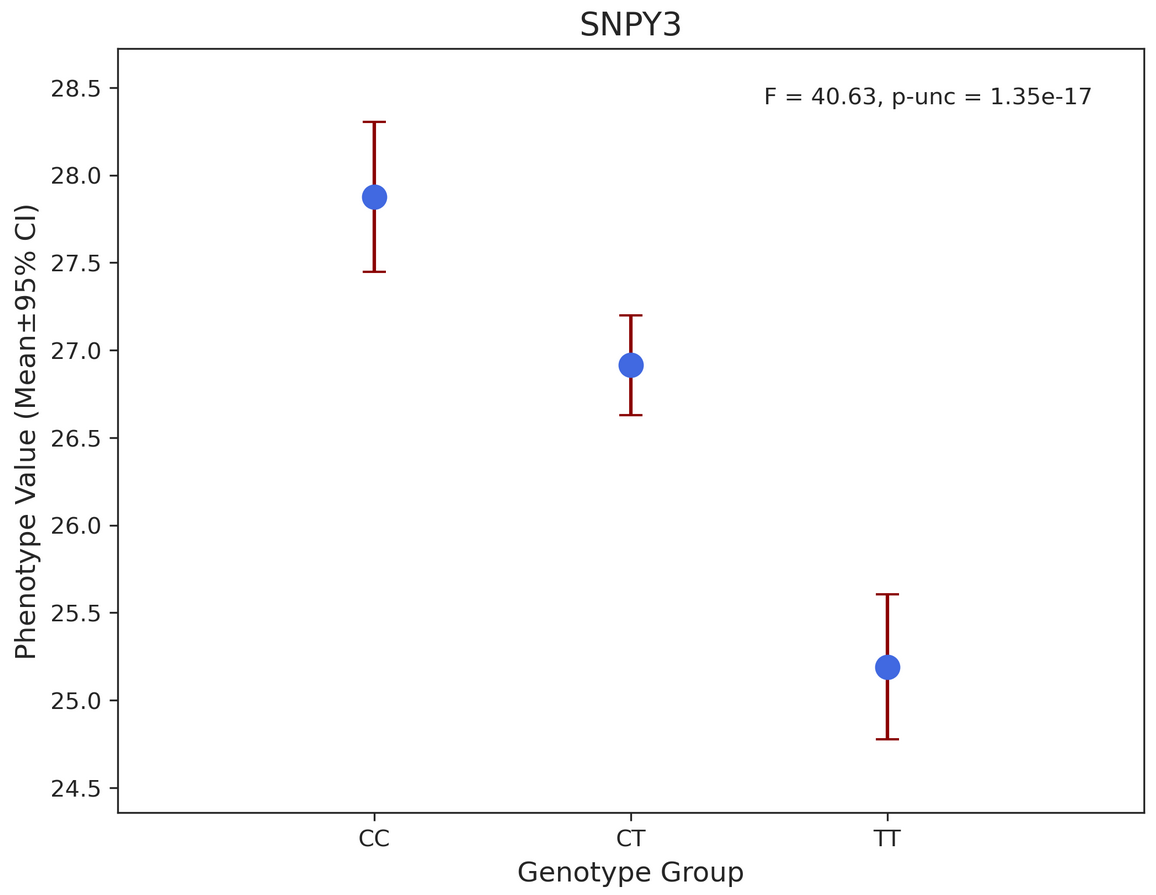
mean(B) : 第二组均值

diff : 组间均值差

p-tukey : 校正后P值,通过Tukey方法调整多重比较的显著性

sig : 显著性标记 p < 0.01（高度显著\*\*）p < 0.05（显著\*）

**（3） 组间均值比较图**



**图1：组间均值比较**

X轴：分组类别；Y轴：表型值均值；误差线：95%置信区间；图中标注F统计量和p值

### 分子标记SNPY4

**（1） ANOVA分析**：单因素方差分析是一种用于比较三个或更多组的均值是否存在显著差异的统计方法，适用于单一分类变量（自变量）对连续型因变量的影响分析。

**表1：单因素方差分析结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | ddof1 | ddof2 | F | p-unc |
| group | 2 | 873 | 4.168e+01 | 5.134e-18 |

Source :变异来源 : "group":（组间变异）,"Residual"或"Error":(未被解释的随机变异),"Total":(整体变异)

ddof1 :组间自由度: 自由度=分组数-1 : 值为2 代表3个分组

ddof2 :组内自由度: 自由度=组内自由度+组间自由度+1 ：值为845 表示总样本量为845,反映独立信息量

F :F统计量 : 组间组内方差比 : 值为58,则表示组间方差是组内方差的58倍

p-unc :未校正p值 : 结果是偶然发生的概率(无效假设)，P值远小于常规阈值0.05 : 拒绝假设

**（2） Tukey HSD**（Honestly Significant Difference，真实显著差异）检验是用于多重比较的统计方法，专门解决ANOVA（方差分析）后多组间两两比较的假阳性问题。

**表2：多重比较检验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A | B | mean(A) | mean(B) | diff | p-tukey | sig |
| CC | GC | 2.514e+01 | 2.694e+01 | -1.801e+00 | 1.000e-03 | \*\* |
| CC | GG | 2.514e+01 | 2.786e+01 | -2.716e+00 | 1.000e-03 | \*\* |
| GC | GG | 2.694e+01 | 2.786e+01 | -9.157e-01 | 1.145e-03 | \*\* |

A : 第一个比较组名

B : 第二个比较组名,AB先后顺序无影响

mean(A) : 第一组均值

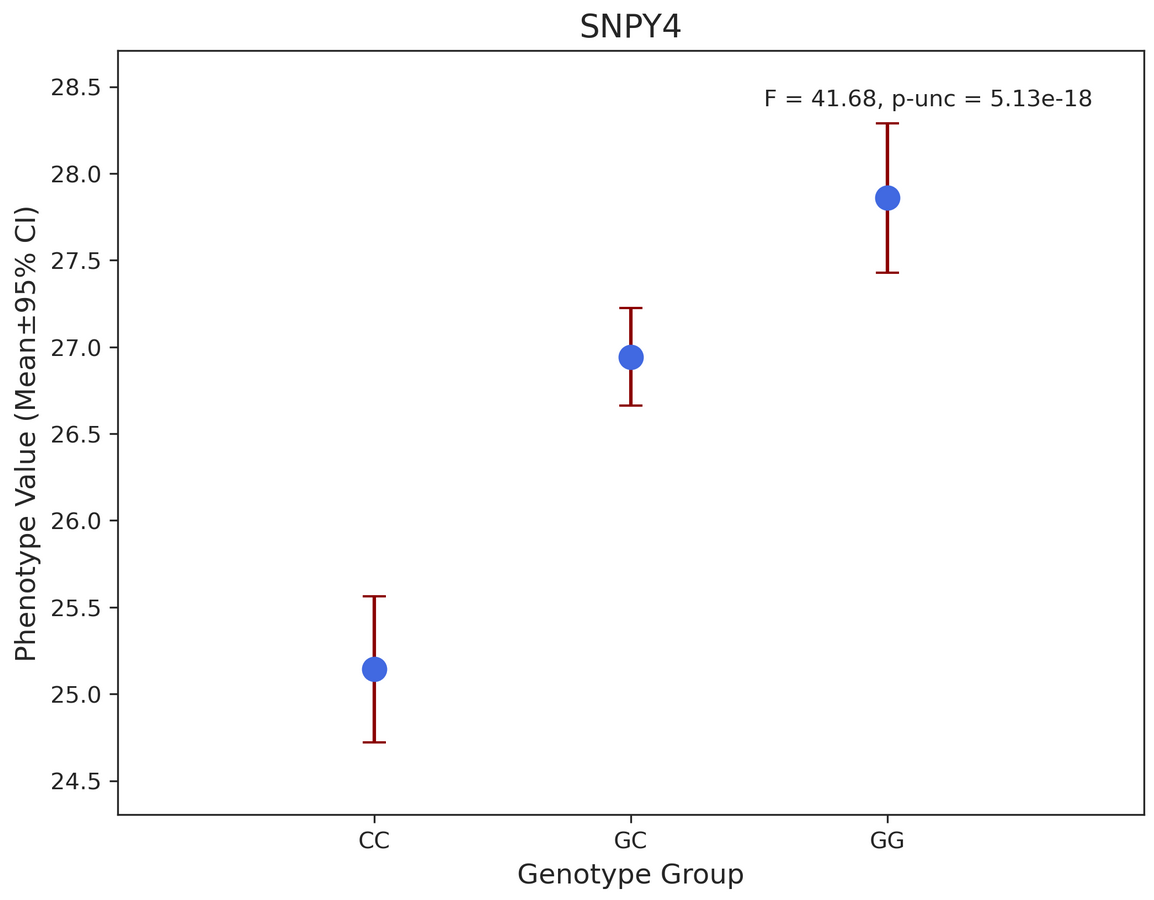
mean(B) : 第二组均值

diff : 组间均值差

p-tukey : 校正后P值,通过Tukey方法调整多重比较的显著性

sig : 显著性标记 p < 0.01（高度显著\*\*）p < 0.05（显著\*）

**（3） 组间均值比较图**



**图1：组间均值比较**

X轴：分组类别；Y轴：表型值均值；误差线：95%置信区间；图中标注F统计量和p值

### 分子标记SNPY5

**（1） ANOVA分析**：单因素方差分析是一种用于比较三个或更多组的均值是否存在显著差异的统计方法，适用于单一分类变量（自变量）对连续型因变量的影响分析。

**表1：单因素方差分析结果**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | ddof1 | ddof2 | F | p-unc |
| group | 2 | 870 | 4.171e+01 | 5.030e-18 |

Source :变异来源 : "group":（组间变异）,"Residual"或"Error":(未被解释的随机变异),"Total":(整体变异)

ddof1 :组间自由度: 自由度=分组数-1 : 值为2 代表3个分组

ddof2 :组内自由度: 自由度=组内自由度+组间自由度+1 ：值为845 表示总样本量为845,反映独立信息量

F :F统计量 : 组间组内方差比 : 值为58,则表示组间方差是组内方差的58倍

p-unc :未校正p值 : 结果是偶然发生的概率(无效假设)，P值远小于常规阈值0.05 : 拒绝假设

**（2） Tukey HSD**（Honestly Significant Difference，真实显著差异）检验是用于多重比较的统计方法，专门解决ANOVA（方差分析）后多组间两两比较的假阳性问题。

**表2：多重比较检验结果**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A | B | mean(A) | mean(B) | diff | p-tukey | sig |
| AA | GA | 2.513e+01 | 2.694e+01 | -1.814e+00 | 1.000e-03 | \*\* |
| AA | GG | 2.513e+01 | 2.786e+01 | -2.726e+00 | 1.000e-03 | \*\* |
| GA | GG | 2.694e+01 | 2.786e+01 | -9.119e-01 | 1.215e-03 | \*\* |

A : 第一个比较组名

B : 第二个比较组名,AB先后顺序无影响

mean(A) : 第一组均值

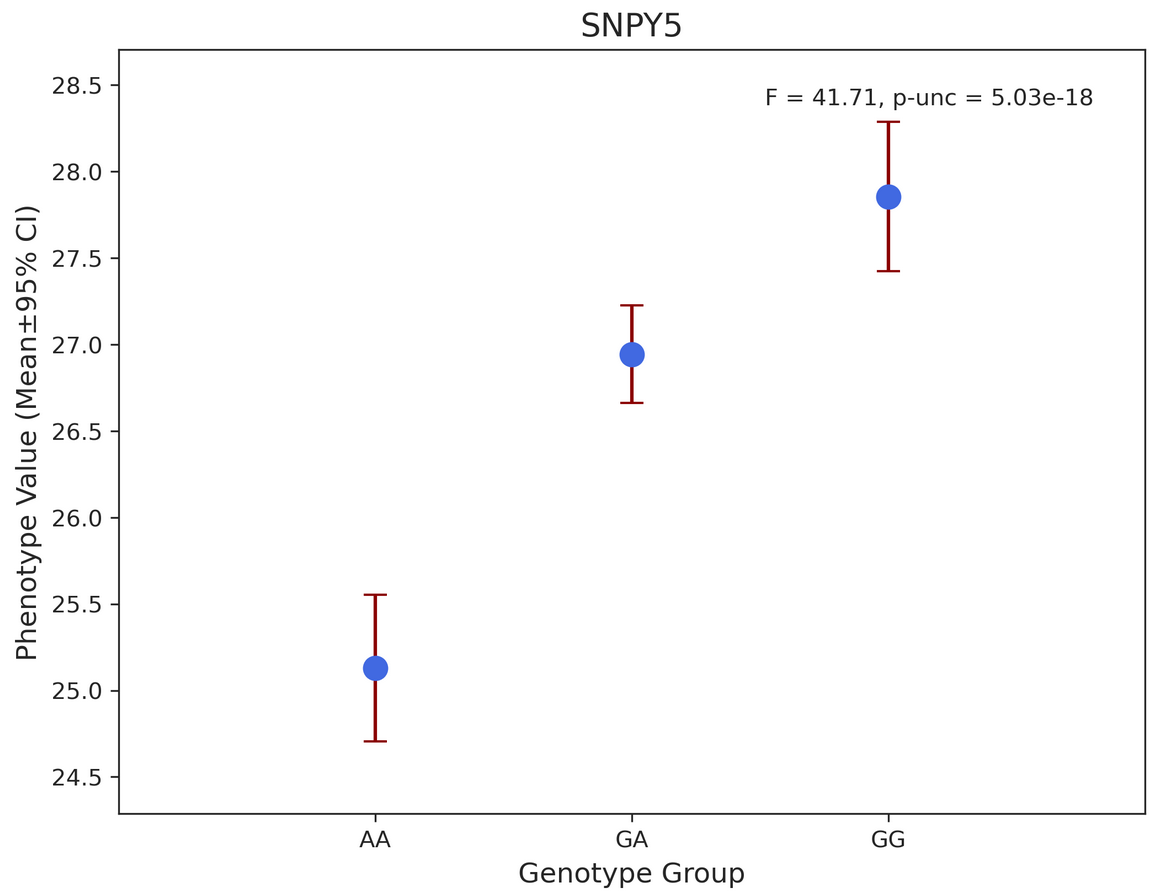
mean(B) : 第二组均值

diff : 组间均值差

p-tukey : 校正后P值,通过Tukey方法调整多重比较的显著性

sig : 显著性标记 p < 0.01（高度显著\*\*）p < 0.05（显著\*）

**（3） 组间均值比较图**



**图1：组间均值比较**

X轴：分组类别；Y轴：表型值均值；误差线：95%置信区间；图中标注F统计量和p值