**实验内容：水稻DUF1645蛋白家族生物学信息分析**

1. **材料与方法**
2. **水稻DUF1645家族成员鉴定及系统进化分析（3.2-3.4）**

从EnsemblPlant数据库（http://plants.ensembl.org/index.html）获得水稻的基因组以及gff3注释文件，通过TBtools软件翻译基因组序列获得水稻蛋白序列。上传DUF1645家族成员的氨基酸序列至NCBI网站（https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi），利用该网站的Web CD-Search Tool功能预测DUF1645家族蛋白序列的保守结构域，通过TBtools的Visualize NCBI CDD Domain Pattern功能进行domain分析的可视化处理。蛋白命名沿用 Yanchun Cui等人的文章[1]。

使用MEGA 7.0软件的Neighbor-Joining（NJ）算法（重复次数为1000来估计分组的置信度）比对DUF1645家族成员的蛋白质序列从而构建其系统进化树，分支点旁边的数字表示支持每个分支的置信度（0-100），数值越大，可信度越高。使用Evolview网站（https://www.evolgenius.info/evolview/#/）及Microsoft Office PowerPoint软件美化系统进化树。OsSGL2使用红色五角星标注。

1. **水稻DUF1645家族基因染色体定位分析（3.15-3.17）**

从EnsemblPlant数据库（http://plants.ensembl.org/index.html）获得水稻的基因组以及gff3注释文件，使用TBtools软件的Gene Location Visualize from GTF/GFF功能分析基因家族成员染色体定位，并使用Microsoft Office PowerPoint美化图片。

1. **蛋白质二、三级结构及信号肽预测（3.24-3.27）**

上传OsSGL2的蛋白质序列到Psipred网站，通过该网站的Protein Structure Prediction工具（http://bioinf.cs.ucl.ac.uk/psipred/&uuid=8f5da9f4-ce07-11ed-bce6-00163e100d53）进行蛋白质二级结构预测。

上传OsSL2的蛋白质序列到SWISS-MODEL网站（https://swissmodel.expasy.org/）进行蛋白质三级预测。

上传OsSGL2的蛋白质序列到SignalP-5.0网站（https://services.healthtech.dtu.dk/services/SignalP-5.0/）进行信号肽预测。

1. **蛋白互作分析（3.6-3.7）**

在STRING数据库（https://cn.string-db.org/）中输入OsSGL2蛋白质序列以得到与其有相互作用的蛋白质的预测结果，选取可信度最高的10个可能与OsSGL2有互作的蛋白进行分析。

1. **水稻DUF1645家族蛋白质理化性质分析（3.19-3.20）**

在TBtools软件的Protein Paramter Calc功能中输入DUF1645家族成员的蛋白质序列，将得到的结果进行整理。

1. **结果分析**

**1. 水稻DUF1645基因家族的系统进化分析（图1）**

水稻基因组中共鉴定出14个DUF1645基因家族成员。系统进化分析表明，DUF1645家族蛋白可分为3个亚家族。其中OsSGL3和OsSGL7关系最为密切。

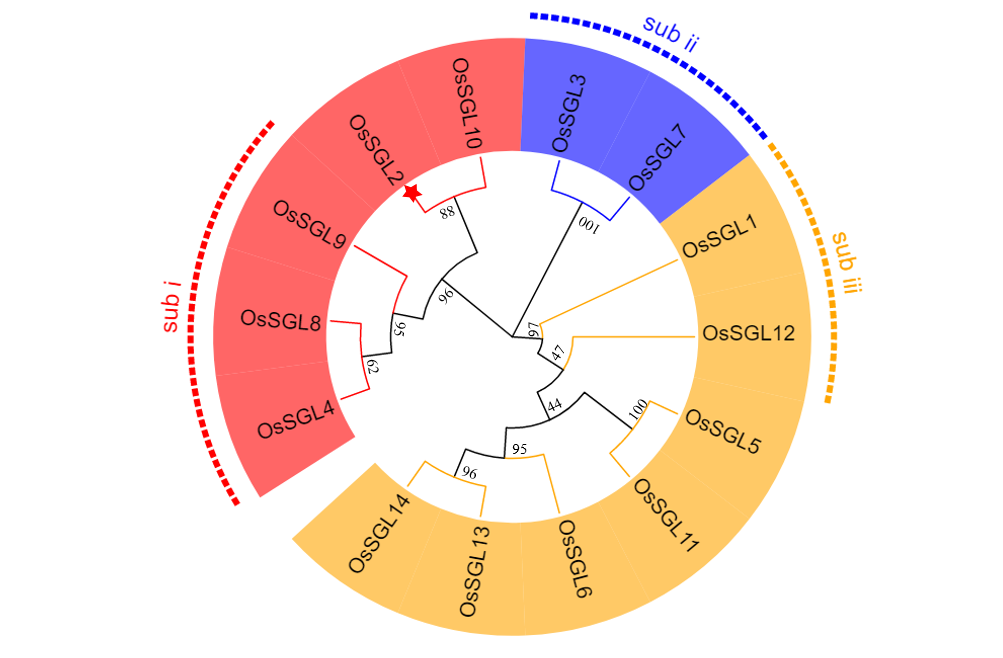


图 1 DUF1645家族基因进化树

**2. 水稻DUF1645基因家族的染色体定位分析（图2）**

*OsSGL2*基因定位于Chr 1，与定位于Chr 5上的*OsSGL10*基因在系统进化上关系最为密切。

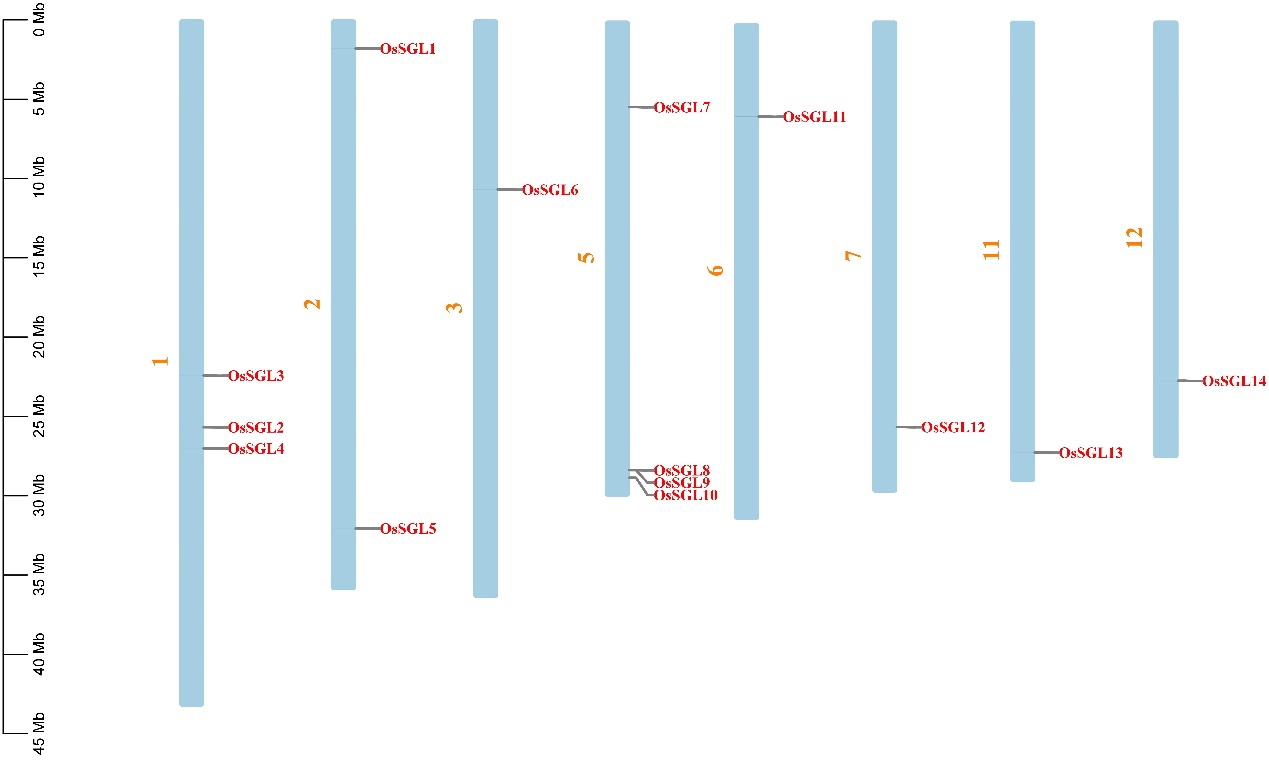


图 2 DUF1645家族基因染色体定位

**3. OsSGL2蛋白质的二、三级结构预测及信号肽预测**

图3为OsSGL2蛋白二级结构预测结果；图4为OsSGL2蛋白三级结构预测结果；图5为OsSGL2蛋白质信号肽预测结果。

由图3可知，OsSGL2蛋白质二级结构拥有7个α螺旋与3个β折叠，α螺旋数目比较多说明其稳定性比较强，β折叠的数目比较少说明其亲水性一般。

由图4可知，OsSGL2蛋白质三级结构整体较为简单，是一个比较对称的结构，能够看到较多的α螺旋与其二级结构的特征相似。

由图5的预测结果可以看出，OsSGL2蛋白上具有信号肽的可能性非常小。

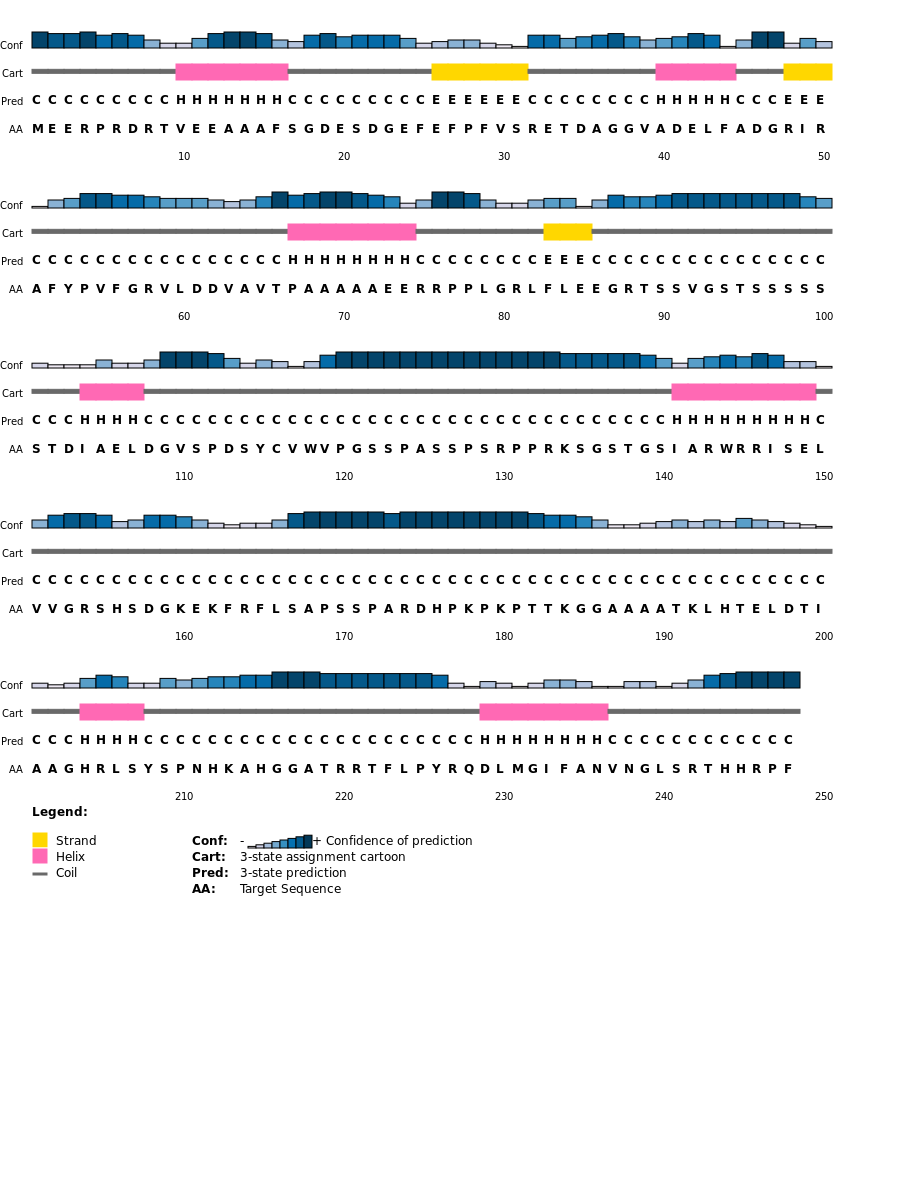


图 3 OsSGL2蛋白质二级结构预测

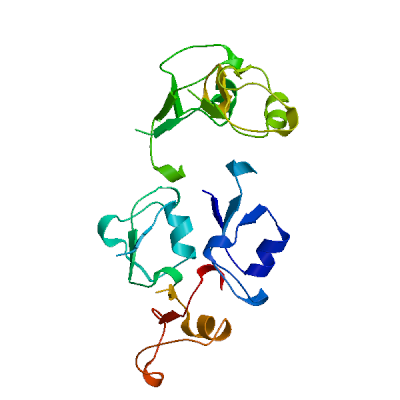


图 4 OsSGL2蛋白质三级结构预测

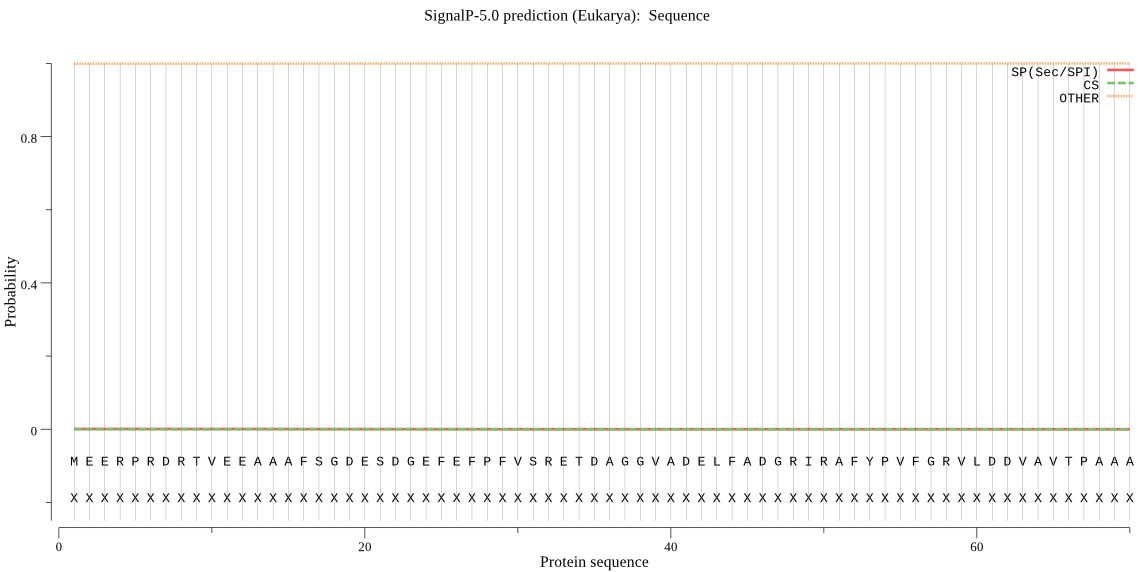


图 5 OsSGL2蛋白质信号肽预测

**4. 互作蛋白分析（图6）**

根据差异基因的功能注释结构可以挑选出OsSGL5、OsSGL6、OsSGL13、LOC\_Os02g56480、LOC\_Os01g08814五个蛋白，它们在表达时同时显示出了高度的相关性，并在子模块分析中也显示出高度的相关性，说明OsSGL2蛋白可能在行使功能时需要与上述五个蛋白共同发挥作用。

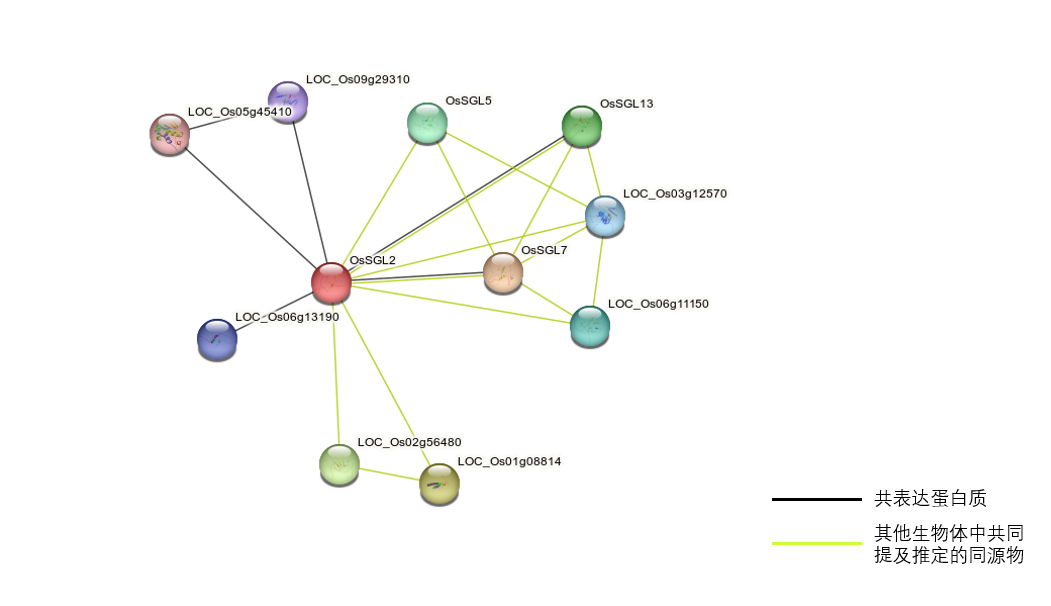


图 6 OsSGL2蛋白互作分析

**5. 水稻DUF1645家族蛋白质理化性质（表1）**

理化性质的分析结果如表1所示，由表1可得OsSGL1、OsSGL6、OsSGL7、OsSGL11与OsSGL2的染色体位置、氨基酸数量、及分子量等性质相差较大，OsSGL2与OsSGL4、OsSGL8的性质最为相近，说明OsSGL2与其家族其他蛋白质在功能上具有一定的差异。

表 1 DUF1645家族蛋白理化性质分析

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| LOC名称 | Os名称 | 重命名 | 染色体位置 | 氨基酸数量 | 分子量 | 理论pl | 不稳定性指数 | 脂肪指数 | 水文性总平均值 |
| LOC ID | Os ID | Rename | Chromosome position | Number of Amino Acid | Molecular Weight | Theoretical pl | Instability Index | Aliphatic Index | Grand Average of Hydropathicity |
| *LOC\_Os01g39760* | *None* | *OsSGL3* | 22417587 - 22418036 | 149 | 15956.69 | 8.62 | 59.79 | 48.52 | -0.795 |
| *LOC\_Os01g45250* | *Os01g0639600* | *OsSGL2* | 25681570 - 25683076 | 248 | 26724.56 | 7.95 | 57.85 | 58.71 | -0.597 |
| *LOC\_Os01g47270* | *None* | *OsSGL4* | 27014866 - 27014072 | 264 | 28519.31 | 5.54 | 84.93 | 51.36 | -0.887 |
| *LOC\_Os02g04130* | *Os02g0134200* | *OsSGL1* | 1800811 - 1799733 | 255 | 26725.75 | 9.35 | 62.63 | 57.29 | -0.5 |
| *LOC\_Os02g52380* | *Os02g0761300* | *OsSGL5* | 32060289 - 32058667 | 430 | 44585.95 | 8.8 | 47.81 | 57.6 | -0.433 |
| *LOC\_Os03g19090* | *Os03g0303100* | *OsSGL6* | 10692291 - 10693924 | 356 | 37144 | 10.04 | 77.73 | 56.07 | -0.602 |
| *LOC\_Os05g09590* | *None* | *OsSGL7* | 5417150 - 5417753 | 193 | 20801.78 | 11.48 | 83.4 | 66.79 | -0.784 |
| *LOC\_Os05g49350* | *Os05g0568600* | *OsSGL8* | 28307054 - 28308078 | 231 | 24054.59 | 5.96 | 64.27 | 64.24 | -0.372 |
| *LOC\_Os05g49370* | *Os05g0568800* | *OsSGL9* | 28310905 - 28311934 | 229 | 24150.07 | 9 | 74.11 | 57.21 | -0.507 |
| *LOC\_Os05g50230* | *Os05g0578100* | *OsSGL10* | 28788169 - 28786911 | 306 | 31933.1 | 5.62 | 53.31 | 53.73 | -0.585 |
| *LOC\_Os06g11150* | *Os06g0215100* | *OsSGL11* | 5847981 - 5849957 | 415 | 43892.9 | 9.51 | 63.08 | 52.87 | -0.697 |
| *LOC\_Os07g42740* | *Os07g0619500* | *OsSGL12* | 25591228 - 25592810 | 329 | 34278.08 | 10.71 | 65.33 | 51.85 | -0.528 |
| *LOC\_Os11g44920* | *Os11g0672900* | *OsSGL13* | 27191011 - 27193021 | 340 | 35669.43 | 10.21 | 89.93 | 53.41 | -0.612 |
| *None* | *Os12g0556900* | *OsSGL14* | 22670886 - 22671923 | 345 | 36290.23 | 11.14 | 94.34 | 48.09 | -0.644 |

1. **实验进度**

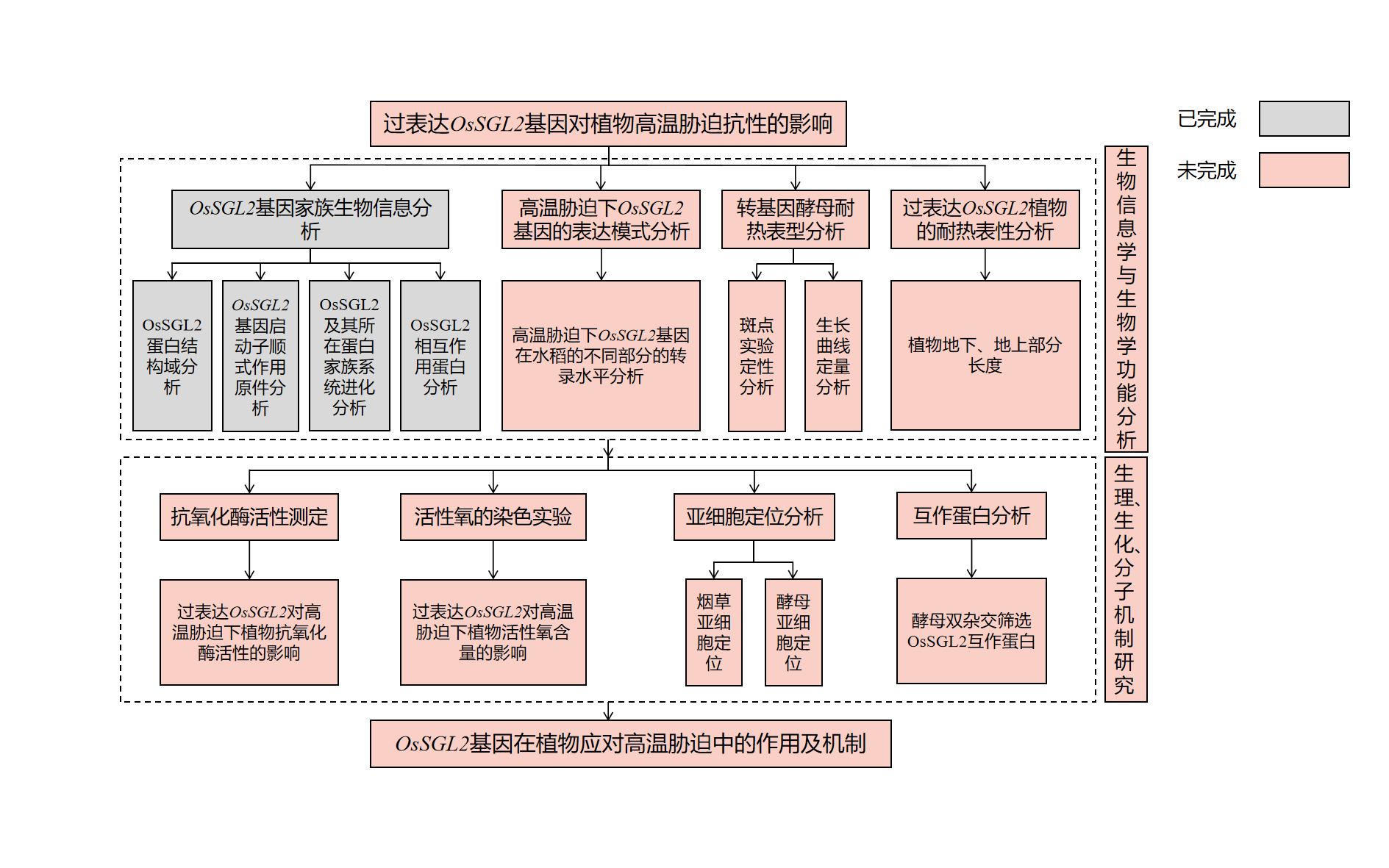


图 7 实验进展图

**参考文献**

1. Cui, Y., Wang, M., Zhou, H., Li, M., Huang, L., Yin, X., . . . Xu, G. (2016). OsSGL, a Novel DUF1645 Domain-Containing Protein, Confers Enhanced Drought Tolerance in Transgenic Rice and Arabidopsis. Front Plant Sci, 7, 2001. doi:10.3389/fpls.2016.02001