# “农业健康与环境”组学大数据整合生物信息学研讨会

# Integrative Bioinformatics Workshop on “Agriculture, Health and Environment”

2017 年 8 月 4-7 日

内蒙古·通辽

# 会议手册

# 论文集

# Proceedings

主办单位

内蒙古民族大学

浙江省生物信息学学会

协办单位

浙江大学、内蒙古科技大学、内蒙古大学、内蒙古农业大学、赤峰学院、内蒙古师范大学

国家烟草基因研究中心、浙江中医药大学、湖南农业大学、江苏省现代作物生产协同创新中心

内蒙古自治区高校蓖麻产业工程技术研究中心及蓖麻育种重点实验室、中国生物工程学会计算与生信专委会

# 会 议 介 绍

## 一、会议目的

为了进一步繁荣我国南北学术交流，讨论“一带一路”的战略机遇，促进生物信息学发展，在自治区成立70周年之际，内蒙古民族大学携手浙江省生物信息学学会，并联合自治区院校及其他科研单位,于2017年8月4-7日在通辽市召开“农业健康与环境”组学大数据整合生物信息学研讨会。

## 二、会议内容

会议就“生物信息学与组学大数据的整合发展、趋势与应用”交流整合生物信息学研究方法、技术与应用的新进展；交流整合生物信息学数据资源平台建设、管理方面的经验；加强科学研究、学术培养、大数据应用方面的合作。会议议题涉及：草原优质基因蛋白资源的发掘；蒙医药生物资源及分子机理研究；草原动植物组学数据的挖掘与整合；最新高通量测序技术与数据分析；比较组学与分子进化；组学大数据的计算问题；基于生物信息学的分子诊断与精准医学；生物信息学与草原生物种群研究；生物资源的收集、保存和发掘利用；系统生物学相关研究进展；生物大数据存储与可视化等方向。

## 三、会议组织

会议主席：陈永胜、陈铭

组织委员会：裴志利、蔡禄、李前忠、李宏、李国婧、范体贵、毕俊怀、曹培健、温成平、易永图、吴玉峰、陈禹保

地方组织委员会：奥·乌力吉、白海花、布仁巴图、姜静清、马德慧、马玉露、韩国志、王思珍、王伟平、温泽、徐程、姚玉华、袁哲明、赵宏林、郑根昌

会议秘书：陈宇杰、沈小仙、陈琦

## 四、会议单位

会议主办：内蒙古民族大学、浙江省生物信息学学会

协办单位：浙江大学、内蒙古科技大学、内蒙古大学、内蒙古农业大学、赤峰学院、内蒙古师范大学、中国烟草基因研究中心、浙江中医药大学医学院、湖南农业大学植保学院、江苏省现代作物生产协同创新中心、内蒙古自治区高校蓖麻产业工程技术研究中心及蓖麻育种重点实验室、中国生物工程学会计算与生信专委会

# “农业健康与环境”组学大数据
# 整合生物信息学研讨会

## 会 议 日 程

| 时间 | | 内 容 | 地点 |
|---|---|---|---|
| 8月4日14:00-22:00 | | **报到** | 通辽市荣达宾馆大堂 |
| 8月5日7:30-9:00 | | | 内蒙古民族大学图书馆一楼大报告厅门厅 |
| 上午 | 9:00-9:10 | **开幕式**<br>内蒙古民族大学校领导及嘉宾致辞 | 图书馆一楼大报告厅<br>主持人：裴志利 教授 |
| | 9:10-9:30 | **集体合影、茶歇** | 图书馆前台阶 |
| | 9:30-10:00 | 报告人：蔡禄 教授（内蒙古科技大学）<br>报告题目：肝细胞癌基因和 miRNA 共表达网络的构建与分析 | 图书馆一楼大报告厅<br>主持人：<br>陈铭 教授<br>黄凤兰 教授 |
| | 10:00-10:30 | 报告人：马闯 教授（西北农林科技大学）<br>报告题目：Systems approach to understanding abiotic stress responses in plants | |
| | 10:30-11:00 | 报告人：李宏 教授（内蒙古大学）<br>报告题目：Nucleosome distributions around transcription start site and their relation to gene expression in yeast | |
| | 11:00-11:30 | 报告人：樊龙江 教授（浙江大学）<br>报告题目：基于人支气管上皮细胞及其组学大数据为基础的烟气风险评价研究 | |
| | 11:30-12:00 | 报告人：李国婧 教授 （内蒙古农业大学）<br>报告题目：柠条干旱胁迫转录组数据分析以及基因功能验证 | |
| 中午 | 12:00-13:30 | **午餐**（自助） | 荣达宾馆自助餐厅 |

**8 月 5 日下午大会报告 （20 分钟/报告人）**

| | | | |
|---|---|---|---|
| 下午 | 14:00-14:20 | 报告人：杨效曾 研究员（北京市农林科学院）<br>报告题目：综合分析和挖掘23种植物中 microRNA 基因的功能元件 | 图书馆一楼大报告厅<br>主持人：<br>徐建红 教授<br>邢永强 副教授 |
| | 14:20-14:40 | 报告人：李晓兰 副教授（赤峰学院）<br>报告题目：线虫对蒙东地区玉米田土壤健康的指示作用研究 | |
| | 14:40-15:00 | 报告人：白海花 教授 （内蒙古民族大学）<br>报告题目：蒙古人全基因组序列图谱构建与东亚人群的进化扩散研究 | |
| | 15:00-15:20 | 报告人：曹培健 副研究员（国家烟草基因研究中心）<br>报告题目：烟草低温胁迫下的转录组和代谢组整合分析 | |
| | 15:20-15:40 | **茶 歇** | |
| | 15:40-16:00 | 报告人：李婧 副教授（上海交通大学）<br>报告题目：生物网络与疾病关键肽段 / 蛋白质的鉴定 | 图书馆一楼大报告厅<br>主持人：<br>陈禹保 教授<br>张文广 教授 |
| | 16:00-16:20 | 报告人：刘金定 副教授（南京农业大学）<br>报告题目：EGAS：一个通用一体化真核基因组注释分析平台 | |
| | 16:20-16:40 | 报告人：许正浩 副研究员（浙江中医药大学）<br>报告题目：雷公藤联合甲氨蝶呤治疗类风湿性关节炎的 meta 分析及生物信息学研究 | |
| | 16:40-17:00 | 报告人：洪艳云 讲师（湖南农业大学）<br>报告题目：非损伤技术分析中国野骆驼群体遗传结构及人口历史动态 | |
| | 17:00-18:00 | **参观内蒙古民族大学博物馆** | |
| | 18:00-19:00 | **晚餐**（自助） | 荣达宾馆自助餐厅 |
| | 19:30-21:00 | **民族声乐欣赏** | 民大音乐厅<br>主持人：马研 |

**8 月 6 日上午大会交流（20 分钟/报告人）**

| 时间 | | 内 容 | 地点 |
|---|---|---|---|
| 上午 | 8:30-8:50 | 报告人：刘贵明 研究员（北京市农林科学院）<br>报告题目：农作物参考基因组的组装和基于高通量测序的表观调控技术及应用 | 图书馆一楼大报告厅<br>主持人：<br>吴玉峰 教授<br>刘鹏 教授 |
| | 8:50-9:10 | 报告人：杨燕 教授（内蒙古农业大学）<br>报告题目：Tamyb10等位基因与小麦籽粒的休眠特性相关 | |
| | 9:10-9:30 | 报告人：相吉山 副研究员（赤峰学院）<br>报告题目：谷子穗发育调控基因（Loose Panicle 1）图位克隆与功能分析 | |
| | 9:30-9:50 | 报告人：金静静 工程师（国家烟草基因研究中心）<br>报告题目：基因型和环境对烟草代谢产物的影响研究 | |
| | 9:50-10:10 | **茶 歇** | |
| | 10:10-10:30 | 报告人：黄凤兰 教授（内蒙古民族大学）<br>报告题目：蓖麻研究工作进展 | 图书馆一楼大报告厅<br>主持人：<br>王瑞刚 教授<br>易永图 教授 |
| | 10:30-10:50 | 报告人：申玉华 副教授（赤峰学院）<br>报告题目：紫花苜蓿逆境响应转录因子 NAC 基因的分离及其功能的研究 | |
| | 10:50-11:10 | 报告人：王晟 讲师（杭州电子科技大学）<br>报告题目：Next generation sequencing of DNA pools | |
| | 11:10-11:30 | 报告人：郭树春（内蒙古大学；内蒙古农牧业科学院）<br>报告题目：Transcriptome sequencing analysis of sunflower responsive to *Verticillium dahliae* infection and mining of resistance-related genes | |
| 中午 | 12:00-13:30 | **午餐**（自助） | 荣达宾馆自助餐厅 |

**8 月 6 日下午大会交流及闭幕式（20 分钟/报告人）**

| | | | |
|---|---|---|---|
| 下午 | 14:00-14:20 | 报告人：黄琳/温成平 教授（浙江中医药大学）<br>报告题目：TCMID 2.0: a comprehensive resource for TCM | 图书馆一楼大报告厅<br>主持人：<br>田迅 教授<br>白有煌 副研究员 |
| | 14:20-14:40 | 报告人：张艾明 讲师（赤峰学院）<br>报告题目：磷在农田土壤中的迁移转化规律及其对农业环境的影响 | |
| | 14:40-15:00 | 报告人：王月 讲师（福建农林大学）<br>报告题目：Genome-wide identification and characterization of putative lncRNAs in the diamondback moth, *Plutella xylostella* (L.) | |
| | 15:00-15:20 | 报告人：王晶晶（浙江大学）<br>报告题目：植物表观遗传修饰的多层次调控研究 | |
| | 15:20-15:30 | **会议闭幕及茶歇** | 主持人：裴志利 教授 |
| | 15:30-18:00 | **南北乒乓球友谊对抗赛** | 体育馆乒乓球运动场<br>主持人：<br>姚玉华 教授<br>徐程 副教授<br>裁判长：王伟平 教授 |
| | 15:30-17:00 | 分区选拔赛，南方队、北方队各选出5名队员 | |
| | 17:00-17:50 | 南北两队各5名选手进行抽签对抗赛，5场3胜制 | |
| | 17:50-18:00 | **颁奖仪式** | |
| | 18:30- | **晚餐**（蒙餐） | 荣达宾馆 |

**8 月 7 日生态环境考察、自由交流**

# “农业健康与环境”组学大数据整合生物信息学研讨会

会议论文摘要

# 目 录

# 肝细胞癌基因和 miRNA 共表达网络的构建与分析

斯阳明，邢永强，蔡禄*

内蒙古科技大学生命科学与技术学院，包头，014010

*E-mail：nmcailu@163.com

**摘要：**肝癌是发生于肝脏组织中的一种常见恶性肿瘤。肝癌具有初期隐匿且临床症状不典型，中后期发展迅速的特点，其发生率和死亡率近年来不断升高，严重威胁着人类的生命和健康。对肝癌发生发展过程中的关键分子标记物的研究，对临床诊断具有重大的意义。TCGA (The Cancer Genome Atlas)项目，即美国政府发起的癌症和肿瘤基因图谱计划，为系统分析癌症组织提供了海量的实验数据。本工作使用 edgeR 工具分析取自 TCGA-LIHC (Liver Hepatocellular Carcinoma)项目的肝细胞癌组织和正常组织基因和 miRNA 表达谱数据，得到肝细胞癌和正常组织的差异表达基因和 miRNA。使用 WGCNA 软件包对得到的差异表达基因和 miRNA 进行共表达网络分析，构建了 27 个与细胞周期、防御反应、甲基化等 GO 分类密切相关的网络模块；正常样本和肝细胞癌样本网络模块的保守性分析发现正常样本和肝细胞癌样本的共表达模式差异显著。最后探索了构建的共表达网络模块与肝细胞癌患者临床信息的关联，发现了 11 个与肝细胞癌患者生存时间、癌症分期等临床信息显著相关的网络模块及 227 个（226 个基因和 1 个 miRNA）与肝细胞癌分期显著相关的网络节点。

**关键词：**肝细胞癌；TCGA；WGCNA；共表达网络；富集分析

# Systems approach to understanding abiotic stress responses in plants

Chuang Ma (马闯)[1,2,3 *]

[1]State Key Laboratory of Crop Stress Biology for Arid Areas, College of Life Sciences, Northwest A&F University, Shaanxi, Yangling 712100, China

[2]Center of Bioinformatics, College of Life Sciences, Northwest A&F University, Shaanxi, Yangling 712100, China

[3]Key Laboratory of Biology and Genetics Improvement of Maize in Arid Area of Northwest Region, Ministry of Agriculture, Northwest A&F University, Shaanxi, Yangling 712100, China

*E-mail：chuangma2006@gmail.com

**Abstract：** As sessile organisms, plants are often subjected to various abiotic stresses (e.g. drought, salinity, extreme temperatures, and oxidative stress) during their lifecycle and they have evolved different mechanisms to combat these stresses [1]. At the molecular level, one of the most immediate responses to stress is the extensive reprogramming of temporal and spatial transcription, resulting in the concurrent rewiring of transcriptional regulation interactions for a large set of genes [2]. To understand how the effect of abiotic stresses on the gene network, we first introduced a Gini-based algorithm, Gini correlation coefficient (GCC), to accurately infer transcriptional regulation relationships based on correlation analysis of gene expression patterns [3]. On the basis of GCC, we then developed systems approaches to thoughtfully investigate abiotic stress responses in model plants and crops [4-6]. By integrating prior knowledge about stress-responsive mechanisms in the model plant *Arabidopsis thaliana*, we developed a machine learning system called mlDNA, to explore the differences between two GCC-based gene networks built using stress-responsive transcriptomes [4, 5]. Experimental results demonstrated that mlDNA has the substantially outperformed traditional statistical testing–based differential expression analysis at identifying stress-related genes, with markedly improved prediction accuracy. For the stress-responsive transcriptomes in crop species such as *Zea mays* (maize), we built a GCC-based gene network from a spatio-temporal transcriptomic map of the drought stress response in maize and characterized hub genes associated with duplication events, selection, and regulatory networks [6]. This integrative strategy provided a new way to identify key transcription factors, which may function as a crosstalk mechanism between drought stress and developmental signaling pathways in maize.

**Keywords：** Abiotic stress; Gini, Machine learning; Network; Transcriptional regulation.

## References

1. Zhu JK. Abiotic stress signaling and responses in plants. Cell, 2016, 167: 313-324.
2. Jogaiah S, Govind SR and Tran LS. Systems biology-based approaches toward understanding drought tolerance in food crops. Crit. Rev. Biotechnol., 2013, 33: 23–39.
3. Ma C and Wang XF. Application of the Gini correlation coefficient to infer regulatory relationships in transcriptiome analsyis.Plant Physiology, 2012, 160: 192-203
4. Ma C and Xin MM, Feldmann KA and Wang XF. Machine learning-based differential network analysis: a study of stress-responsive transcriptomes in Arabidopsis. Plant Cell, 2014, 26:520-537.
5. Ma C, Zhang HH and Wang XF. Machine learning for big data analytics in plants. Trends in Plant Science, 2014, 19(12):798-808.
6. Miao ZY, Hang ZX, Zhang T, Chen SY and Ma C. A systems approach to a spatio-temporal understanding of the drought stress response in maize. Scientific Reports, 2017, 7: 6590.

# Nucleosome distributions around transcription start site and their relation to gene expression in yeast

Hu Meng, Hong Li*, Yan Zheng, Yun Jia, Zhenhua Yang

孟虎，李宏*，郑燕，贾云，杨振华

Laboratory of Theoretical Biophysics, School of Physical Science and Technology, Inner Mongolia University, Hohhot, 010021, China

*E-mail: ndlihong@imu.edu.cn

**Abstract:** Based on the single-base pair resolution map of nucleosome positions in yeast, we used two types of position definitions for +1/-1 nucleosomes to analyze nucleosome distribution patterns on transcription start site (TSS) flanking regions. While the position definition based on TSS, it could not reflect the functional distributions of nucleosomes. Then we re-demarcated +1/-1 nucleosomes by nucleosome free region (NFR) in 6664 yeast protein coding genes. About 2/3 genes have striking NFR (sNFR) and about 1/3 have non-striking NFR (nNFR). In sNFR genes, the length and the order structure of NFR displayed a positive correlation with gene expression levels. In nNFR genes, the position of -1 nucleosome has obvious conservative property and does not correlate with gene expression levels. Further, through the constitutive analysis of +1/-1 nucleosomes sequences in sNFR/nNFR genes, it showed that -1 nucleosome sequences in nNFR genes are different from that in sNFR genes. Two kinds of gene transcriptional regulatory modals appear in our analysis. One is NFR dependence modal and the other is NFR independence modal. In the first mechanism, the transcriptional efficiency is tightly related to the modifiability of -1 nucleosome position or the length of NFR and the order structure intensity of NFR sequence. In the second mechanism, the transcription initiation depends on the -1 nucleosome dispelling and the transcriptional efficiency is regulated by the combination intensity between -1 nucleosome core sequence and histone octamer. This conclusion is similar with the Kubik's recent research. But, it is still controversial whether the nucleosome dispelling mechanism exists, and our theoretical analysis thought that this mechanism does exist in yeast.

**Acknowledgments:** This work was supported by grants from the National Natural Science Foundation of China (No. 31260219), doctor subject Foundation of the Ministry of Education of China (No. 20121501110006).

# 基于人支气管上皮细胞及其组学大数据为基础的烟气风险评价研究

沈一飞，陈曦，樊龙江*

浙江大学作物科学研究所、生物信息学研究所&空气污染与健康研究中心，紫金港校区，杭州 310058

*E-mail: fanlj@zju.edu.cn

**摘要：**吸烟对人体最直接的影响主要集中在对支气管及肺的危害上，科学合理的评价烟气对人体肺部的影响程度，不仅是科学研究的一个重要领域，而且还是对卷烟危害性评判的一个重要依据。所以，构建一套能够真实模拟在吸烟状态下人体支气管（肺）的生理状态的评价体系具有非常重要的科研及实践意义。随着国内烟草企业的国际化程度不断提高和烟草产品的不断改进升级，对减害卷烟的评价不仅仅要看烟气中有害的化学成分是否降低，更需要着重评价烟气对人体细胞影响程度是否有无下降。因此，对传统及新型卷烟危害性的比较评价十分紧迫及重要。本研究建立了一个科学评价不同烟气对人支气管（肺）上皮细胞影响差异的综合评价体系。该体系中包括模仿人肺环境的细胞烟气暴露培养实验平台和利用计算机进行高通量实验数据分析平台等。利用该评价系统，可以更加科学客观的对卷烟等烟气的风险进行评判。该体系以人肺支气管上皮细胞为模型，研究全基因组的基因表达，蛋白质翻译以及最后的代谢产物这三个方面，对受到不同烟气处理的细胞进行全面的分析及评价，最终建立一个基于人类组学大数据科学分析的吸烟风险评价系统。该评价体系为目前国内第一个人体支气管（肺）上皮细胞为模型的烟气危害评价实验平台。相对于以往的传统毒理学评价指标（如小鼠吸入急毒试验、细胞毒性试验、Ames 试验和细胞微核试验），该评价模型更能真实、准确地反应烟气对人体支气管与肺部的影响情况。以电子烟为例，进行了电子烟与传统卷烟烟气对上皮细胞转录组影响的比较分析及其风险评价。

**关键词：**吸烟风险评价；人肺支气管上皮细胞；烟气暴露生物评价平台；转录组；组学大数据

# 柠条干旱胁迫转录组数据分析以及基因功能验证

杨杞，韩晓敏，于秀敏，赵娜，王光霞，万永青，杨飞芸，张燕娜，冯宗琪，李娜，毛铭铢，郑如玉，滑璐玢，刘坤，岳文冉，王瑞刚，李国婧*

内蒙古农业大学生命科学学院，呼和浩特，010018

*Email: liguojing@imau.edu.cn

**摘要：**柠条是柠条锦鸡儿(*Caragana korshinskii*)、中间锦鸡儿(*Caragana intermedia*)和小叶锦鸡儿(*Caragana microphylla*)的统称，属于豆科锦鸡儿属多年生灌木，是干旱草原、荒漠草原地带的优良固沙和水土保持植物，具有很高的生态价值和经济价值。本实验室综合利用二代测序技对干旱处理后 0h、1h、3h 和 12h 的中间锦鸡儿幼苗进行转录组深度测序，结果如下：1.得到 1271 条共差异表达序列（ESTs），其中上调表达 766 条，下调表达 505 条， 对部分 Unigenes 的 qRT-PCR 检测结果证明转录组结果可靠。2.从差异表达 ESTs 中筛选到 404 个 SSR 标记，并从中筛选出多态性高、重复性好的引物 7 对。3.对 NAC、MYB、WRKY、bHLH、AP2 等转录因子家族和 LEA、泛素连接酶、PP2C、黄酮代谢途径酶类、DUF 等功能基因家族成员与豆科植物以及模式植物拟南芥的聚类分析显示，柠条的干旱诱导表达基因与蒺藜苜蓿、鹰嘴豆、百脉根等物种的同源基因更相似。4.以模式植物拟南芥为转化系统，对上述基因家族部分成员的功能分析表明：绝大多数基因参与了对干旱、盐、衰老等的抗性（Xiaomin Han et al, BMC Plant Biology; Qi Yang et al, Molecular Biology Reports ; 杨杞等，林业科学；赵娜，硕士毕业论文；王光霞，博士毕业论文；于秀敏，博士毕业论文）。

**关键词：**柠条；转录组；功能分析；转录因子；基因家族

# 综合分析和挖掘 23 种植物中 microRNA 基因的功能元件

杨效曾*，王颖，李磊

北京市农林科学院，北京，100097

*E-mail: yangxiaozeng@baafs.net.cn

**摘要：**尽管越来越多的植物 microRNA（miRNA）被发现和鉴定，但是现有数据库中的很多物种 miRNA 并不全面。本文通过挖掘由第二代测序（Next generation sequencing-NGS）技术产生的小 RNA 文库，系统地鉴定 23 种植物中的 miRNA 基因。通过分析 miRNA 基因的前体序列、二级结构、成熟序列和启动子，揭示了这些元件和因素在 miRNA 功能进化分化中的角色。

# 线虫对蒙东地区玉米田土壤健康的指示作用研究

李晓兰

赤峰学院生命科学学院/农业科学研究院，内蒙古 赤峰 024000

**摘要：**内蒙古东四盟是中国东北玉米带的重要组成部分，该区域玉米种植面积占内蒙古自治区玉米总种植面积的一半以上。近些年来，玉米的种植包括覆膜和不覆膜两种，其中覆膜面积较大。随着覆膜连作年限的增加，土壤中留存的残膜量也呈现增加趋势，并且残膜的存在影响了土壤结构。本文以土壤线虫为指示生物，研究蒙东地区玉米田多年覆膜连作后残膜的污染状况，以及随着残膜量的增加对土壤线虫群落的影响，通过线虫群落变化与土壤理化性质的关系分析线虫对土壤系统变化的指示作用，进而为蒙东地区土壤的可持续发展的相关研究提供依据。通过研究主要得出以下结论：

（1）连续覆膜 3 年后，土壤中 0-30cm 范围内残存的地膜量已达到 83.3kg/hm$^2$。对土壤含水量、阳离子交换量、有机碳、全磷和土壤质地等土壤理化性质已有显著影响。

（2）连续覆膜 4 年后，土壤中线虫数量显著降低。营养类群中的植物寄生线虫、食真菌线虫和食细菌线虫数量均显著减少，相对丰富度变化不明显。连续覆膜 4 年后，生活史类群中的 cp2 类群数量和相对丰富度均下降，而 cp4 类群却显著上升。

（3）连续覆膜 4 年后，植物寄生线虫中的短体线虫属显著增加，该属数量的增加会引起植物根系的腐烂，组织的坏死，是造成土传病害的原因之一。

从结果可以看出，土壤线虫数量和种类的变化能够指示土壤环境的变化，在本研究中，连续覆膜 4 年种植玉米已经对土壤产生了显著影响，继续连续种植可能会引起病害的发生或减产，该结果对指导覆膜玉米的种植具有重要意义。

**关键词：玉米；残膜；土壤线虫；蒙东地区**

# 蒙古人全基因组序列图谱构建与东亚人群的进化扩散研究

白海花*，吴柒柱，布仁巴图，吉日木图，裴志利，陈铭，陈宇杰，高颖，苏雅拉图，杨立清

内蒙古民族大学，通辽，028043

*E-mail: hh_bai@163.com

**摘要：**本研究对典型的 1 例蒙古族个体进行 de novo 测序和 200 例蒙古族个体进行重测序，共得到约 150 G 的 Reads 总数，13T 的碱基总数，平均测序深度达到 21X。研究对象为自愿入组的新疆-卫拉特蒙古人（40 人），内蒙古布里亚特蒙古人（35 人）、科尔沁蒙古人（40 人）、阿巴嘎旗蒙古人（35 人）和苏尼特蒙古人（16 人）以及蒙古国的乌拉巴托蒙古人（35 人）。

由于测序样本和数据的相对缺乏，中亚/北亚人群的遗传变异及其进化等重要科学问题尚未得到一致的结论。为了解决这个问题，本研究构建了蒙古族全基因组序列图谱，并与千人基因组数据比对分析，确定 1520 万单核苷酸多态性（SNPs），其中 390 万（25.8%）是新发的。本研究发现了芬兰人和蒙古人之间的共同祖先等位基因，并表明了欧洲人与东亚/北亚之间的基因交流路径。此外，本研究中的多条证据支持现代东亚人的祖先可能通过西西伯利亚从北到南移的假设。总之，蒙古人全基因组序列图谱的构建将成为确定中亚/北亚人群的进化路径研究以及个体化医疗等方面提供宝贵的遗传资源。

# 烟草低温胁迫下的转录组和代谢组整合分析

金静静，张慧，张剑锋，刘萍萍，陈霞，李泽锋，许亚龙，卢鹏，曹培健*

国家烟草基因研究中心，中国烟草总公司郑州烟草研究院，郑州，450001

*E-mail：peijiancao@163.com

**摘要：**翠碧一号（CB-1）和 K326 是非常接近的两个烟草品种，但是它们抗低温胁迫能力缺差异很大，K326 比 CB-1 更耐冷。我们利用转录组和代谢组学技术手段，分析了 CB-1 和 K326 在低温处理和无处理下的胁迫响应。CB-1 品种里鉴定了 14590 个差异表达基因，K326 品种里鉴定了 14605 个差异表达基因，同时我们获得了 200 多个差异代谢物。差异表达基因中，50%左右同时在 CB-1 和 K326 中响应，但很多基因在 K326 中响应程度明显高于 CB-1。差异代谢产物中，我们发现了同样的规律，尤其是在一些重要的初生代谢产物上，如氨基酸、有机酸和糖类。同时，我们还发现在 CB-1 和 K326 中能量代谢和激素代谢途径表现出了相反的规律。研究表明，基因转录和代谢产物的差异可能是导致 CB-1 和 K326 抗低温胁迫能力差异的主要原因。

**关键词：**转录组；代谢组；烟草；低温胁迫

# 生物网络与疾病关键肽段／蛋白质的鉴定

徐佳，雷明莉，王博，李婧*

上海交通大学生命科学技术学院生物信息学与生物统计学系，上海，200240

*E-mail: jing.li@sjtu.edu.cn

**摘要：**蛋白质突变、修饰以及差异表达在疾病发生发展以及人类对疾病机理对理解都扮演重要角色。随着基于质谱技术的蛋白质组技术的快速发展，高通量蛋白质组学已经成为复杂样本中蛋白质定性、定量分析的主流方法。那么如何鉴定关键性、驱动性的肽段与蛋白质是关键科学问题，也依然面临巨大挑战。目前已有的相关算法与工具往往忽略了生物细胞内蛋白质相关作用的特性，视每一个蛋白质为独立单位。这里，我们将聚焦生物网络，报告一系列的基于生物网络的统计模型与生物信息学算法，实现疾病关键肽段、蛋白质组的鉴定、注释与风险评估。仿真模拟与实例应用都表明，我们提出的新策略由于考虑了突变、修饰与蛋白表达谱的相关，可以更有效地识别关键的肽段、蛋白质。这些肽段与蛋白质也可以作为潜在的致病因子、药物靶点与预后标志物。

**关键词：**蛋白质变异；翻译后修饰；差异表达；生物网络

# EGAS：一个通用一体化真核基因组注释分析平台

刘金定

南京农业大学信息科技学院，南京，210095

E-mail：liujd@njau.edu.cn

**摘要：**随着高通量测序技术发展，生物学研究中获得基因组序列变得越来越容易，大量真核基因组测序项目得以启动并相继完成。真核基因组注释分析工作具有跨学科、数据容量大、计算复杂性高等特点，因此基因组注释分析工作相对滞后。为此，我们设计并实现了一个通用一体化的在线服务平台（EGAS：Eukaryotic Genome Analysis Server）用于真核基因组注释和分析工作。用户通过 WEB 接口提交任务后，一个自动化的分析处理流程将被创建。该流程不仅涵盖了重复元件、非编码 RNA 和编码基因等常规的基因组结构注释功能，而且还包括上游基因组组装质量评估、下游编码基因功能注释和比较基因组学分析等功能。此外，它还能够产生丰富的图和表帮助研究人员快速了解基因组特征。

**关键词：真核基因组；基因组注释，比较基因组学；Web 服务**

# 雷公藤联合甲氨蝶呤治疗类风湿性关节炎的 meta 分析及生物信息学研究

王晓宇，祖筠筠，黄琳，温成平，许正浩*

浙江中医药大学基础医学院，杭州，310053

*E-mail：xuzhenghao@zcmu.edu.cn

**摘要：**雷公藤联合甲氨蝶呤在中国已经被广泛应用于类风湿性关节炎(Rheumatoid arthritis, RA)的治疗，但是该联合用药策略的安全性、有效性及其机制还不是很清楚。本研究希望利用 meta 分析考察该治疗方法的安全性和有效性，并进一步利用生物信息学初步研究其潜在的作用机制。在中英文数据库(中国知网、维普、万方、Pubmed 和 Web of Science)中全面检索雷公藤联合甲氨蝶呤治疗 RA 的临床文献研究。对文献进行筛选以及相关数据的提取和分析，对雷公藤联合甲氨蝶呤的有效性和安全性进行 Meta 分析研究。同时，从 TCMID 数据库中搜索雷公藤相关作用靶点；从 OMIM、GAD、KEGG 中获取 RA 疾病的相关基因；从 DrugBank 等数据库中获取治疗 RA 的药物相关作用靶点。对所获得的这些基因和靶点进行比较和分析相关信号通路，对雷公藤联合甲氨蝶呤治疗 RA 可能的作用机制进行初步的信息学研究。最终，对符合纳入标准的 6 篇（共计 643 例患者）联合用药的随机对照研究进行 meta 分析发现：（1）疗效方面，联合雷公藤可提高 50%有效率(RR 1.337, 95%CI: 1.188-1.505, P < 0.001)，并缩短晨僵时间、减少疼痛和肿胀关节数、降低红细胞沉降率、降低 C 反应蛋白表达、减少类风湿性因子；（2）安全性方面，合用雷公藤多苷并不增加不良反应发生率 (RR 0.824, 95%CI: 0.635-1.068, P = 0.143)。生物信息学分析进一步发现雷公藤和甲氨蝶呤可能存在 6 个共同的作用靶点，两者可能共同作用于淋巴细胞增殖相关信号通路。此外，雷公藤还可能单独调节免疫系统的其他三条 RA 相关的信号通路（调节单核细胞、TLR1：TLR2 信号通路、干扰素 γ 信号通路）。因此，本研究初步表明雷公藤联合甲氨蝶呤较单独应用甲氨蝶呤可能是一种安全有效的治疗策略，并且生物信息学研究为该用药方式的有效性提供理论证据。雷公藤联合甲氨蝶呤的长期疗效、安全性及其作用机制仍需要今后进一步的研究。

**关键词：**类风湿性关节炎；雷公藤；甲氨蝶呤；meta 分析；生物信息学分析

# 非损伤技术分析中国野骆驼群体遗传结构及人口历史动态

洪艳云[1,2,*]，吴鹏[2]，李迪强[2,*]

[1]湖南农业大学植物保护学院，长沙，410128；

[2]中国林业科学研究院森防森保保所，北京，100091

*E-mail: diqiangli@yahoo.com.cn, hongyanyun@126.com.

**摘要：**野骆驼（*Camelus ferus*）是旧世界真驼属唯一的野生种，数量不足 830 峰，是比大熊猫还珍稀的极度濒危野生保护动物，它们仅存于蒙古国西南和我国西北的库姆塔格沙漠和塔克那玛干沙漠。我们借助 5 对骆驼科特异性微卫星引物，对来自塔克拉玛干沙漠（TK）、库姆塔格沙漠的阿尔金山北麓（AJ）、西湖湿地（XH）和小泉沟（XQG）4 个野骆驼分布区的 1028 份野骆驼粪便样品的个体进行了识别，并对识别个体遗传多样性，群体遗传结构，群体人口历史动态进行了研究。结果显示：1）5 对微卫星引物共扩增得到 101 个多态性基因片段，各位点平均等位基因数(Na)为 14.65，各位点在全部样品中平均有效等位基因数(Ne)为 8.03，各位点在全部样品平均观测杂合度(Ho)是 0.37 之间，各位点在全部样品中平均期望杂合度(He)是 0.87，各位点在全部样品中固定指数是 0.57。2）1028 份粪便样品来自 420 峰野骆驼，Structure 等生物学软件分析发现这 420 峰野骆驼分为两个群体：TK 单独为一个群体（A），而 AJ、XH 和 XQG 为另一个群体（B）。地理群体遗传分化结果显示：TK 与 AJ、XQG、XH、3 个群体分化最明显，与它们的遗传距离依次为 0.359、0.416、0.517。而 AJ、XQG 及 XH 三个地理群体两两之间遗传距离较小，XH 与 AJ 遗传距离最大为 0.246； XH 与 XQG 距离交次之，为 0.139；而 AJ 与 XQG 遗传距离最小，仅为 0.111 方差。方差分析显示（AMOVA），差异的 95%来自群体内和个体间，只有 5%的差异来自群体间。3）野骆驼 历史上曾经经历了两次人口紧缩，最近的一次人口紧缩在生在 1000 年左右。我们的研究结果将为极度濒危的野骆驼保护政策的制定与执行提供理论依据。

**关键词：野骆驼，粪便，非损伤评估，人口历史动态，遗传结构**

# Genetic variation, population structure and history of demography of *Camelus ferus*

Yanyun Hong[1,2], Peng Wu[2], Diqiang Li[2]

[1]College of Plant Protect Science and Technology, Hunan Agricultural University, Changsha, 410128, China; [2]Institute of Forestry Ecology, Environment and Protection, and The Key Laboratory of Forest Ecology and Environment of State Forestry Administration, Chinese Academy of Forestry, Beijing ,100091, China.

**Abstract:** Camelus ferus are the alone survivors of the old world camels. At present, their total number is only about 730–880, less than that of Giant panda (Ailuropoda melanoleuca). They live in Southwestern of Mongolia and Northwestern of China, which are Kum Tagh desert and Taklimakan deser, considered critically endangered by the International Union for Conservation of Nature. In the study,1028 scats were collect from Altyn Mountain (AJ) , Xioqungou (XQG),Xihu wetland (XH) in and around the Kum Tagh desert region, and Taklimakan desert(TK), 5 microsatellites loci were used to analyze the genetic diversity, population genetic structure and history of demography of Camelus ferus. The genetic diversity showed the average number of effective alleles was 8.03, a total of 101 alleles were observed across all the four populations with mean number of alleles per locus as 14.65. The mean observed heterozygosity (Ho) and expected heterozygosity (He) were 0.37 and 0.87. Then we found 420 individuals from 1028 fecal samples. The analysis result of the genetic distance among population indicated that genetic distance between TK and other three population (AJ, XQG and XH) was farthest, which were 0.359、0.416、0.517;and nearest between AJ and XQG（0.111）. Results of analysis of molecular variance showed that the variance among populations was 5%, and the variance among individuals was 95%. The genetic differentiation index indicted that the genetic differentiation degree was very high among individuals with in population. Population structure and degree of admixture were assessed using the program STRUCTURE and the estimated number of clusters in all simulations was K=2. The one was TK group, the other was made up of AJ,XQG and XH. The conclusion validated that geographic block was one of the main factors that affected wild animal genes flow. Bottleneck tests revealed that Camelus ferus experienced a severe population decline (50-fold) during the past 1000 years. The current effective population size (Ne) of Camelus ferus is less than 1000and the ratio of Ne to the census population size is approximately 0.08. The work will help to develop a more accurate protection policy for Camelus ferus and facilitate further research on selection and adaptation of Camelus ferus.

**Keywords:** Camelus ferus, scats, noninvasive assessment, population structure, demographic history.

# 农作物参考基因组的组装和基于高通量测序的表观调控技术及应用

刘贵明

北京农林科学院

E-mail: liugm@big.ac.cn

**摘要：**随着高通量测序技术的不断发展和测序成本逐渐降低，三代测序技术以其独特的长读长优势，可以很好地解决重复序列高、杂合率高的基因组组装。结合 BioNano 光学图谱和 Hi-C 三维基因组技术构建高质量的参考基因组，同时利用长 read 精确分析染色体结构变异(structural variation，SV)。基于 NGS 测序数据的 RNA-seq、ChIP-Seq、ATAC-Seq、BS-seq、HiC-seq、DAP-seq 和 CrY2H-seq 技术，从全基因组水平研究转录组、组蛋白修饰、染色质状态、甲基化、染色体高级结构和 DNA/蛋白质互作，使基因组 SV 和表观遗传研究成果为作物品种改良提供新的科学依据。

# Tamyb10 等位基因与小麦籽粒的休眠特性相关

王瑜，王晓丽，孟建宇，张颖君，何中虎，杨燕*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail：yangyanchutao@imau.edu.cn

**摘要：**小麦籽粒颜色影响面粉的色泽、种子休眠及穗发芽抗性。转录因子*Tamyb10*影响小麦籽粒颜色和休眠特性。本研究在一套具有不同休眠特性的中国面包小麦中发现*Tamyb10*具有不同的等位变异，并且一个功能型的STS标记*Tamyb10D*被开发，该标记可以有效地鉴定小麦籽粒的休眠特性。应用该标记可以扩增出两种PCR片段类型，在休眠性强的材料中可以扩增出 1629-bp 和1178-bp两种片段，而在休眠性弱的材料中只能扩增出1160-bp 的片段。用103份具有不同休眠特性的小麦自然群体对该标记的有效性进行验证，数据统计分析表明在这套材料中该标记*Tamyb10D*与小麦籽粒的休眠特性显著相关（$P<0.001$）。另外用一套重组自交系（中优9507 ×洋小麦）对该标记的有效性进行进一步的验证，中优9507 (1178-bp)和洋小麦(1178-bp 和1629-bp)两种亲本的籽粒分别具有极弱和极强的休眠特性。一般线性模型分析表明*Tamyb10-D1* 的不同等位变异与发芽指数具有显著相关性（$P<0.001$），在两个不同的生态区可以解释13.7%和4.7% 表型变异率。在103份的自然群体中共发现 8种*Tamyb10* 基因型（包括*Tamyb10-A1*、 *Tamyb10-B1* 和 *Tamyb10-D1* 三个位点）分别命名为 aaa、aab、aba、abb、baa、bab、bba 和 bbb，这些不同的基因型与发芽指数值显著相关。

**关键词：等位变异，发芽指数(GI)，*Tamyb10-D1*，休眠**

# 谷子穗发育调控基因（*Loose Panicle* 1）图位克隆与功能分析

相吉山[1,2]，汤沙[2]，李晓兰[1]，王延丽[1]，智慧[2]，贾冠清[2]，刁现民[2,*]

[1]赤峰学院生命科学学院/农业科学研究院，赤峰 024000；

[2]中国农业科学院作物科学研究所，北京 100081

*E-mail：diaoxianmin@caas.cn

**摘要：**谷子（*Setaria italica* (L.) P. Beauv.）属于禾本科（Gramineae）狗尾草属（Setaria Beauv）二倍体（2n=2x=18）植物，具有基因组小（约 515 Mb）、重复序列少、生育期短、繁殖系数高等特点。目前，谷子已完成了全基因组测序，并构建了高密度分子标记连锁图谱和单倍型图谱，正发展成为禾本科新的模式植物，是研究作物抗旱节水、$C_4$ 机理等的理想材料。

谷子属于稀播作物，产量基本靠主茎成穗，穗部性状是决定产量的主要因素。谷子的穗属于穗状圆锥花序，其分枝数多、颖花数多、花器官小、穗粒数多。目前，关于调控谷子穗发育基因的研究鲜有报道。本研究以 EMS 诱变谷子全基因组测序品种豫谷 1 号（Yugu1）所获得的穗发育异常突变体 *loose panicle 1*（*lp1*）为对象，通过表型调查，研究穗发育异常等的形态特征；利用 BSA、MutMap 法图位克隆目的基因；通过亚细胞定位、表达量分析，探讨目的基因的表达特征，解析目的基因影响谷子穗发育的功能。主要结果如下：

与野生型相比，*lp1* 的植株主茎变矮，主穗穗长增长、第一级分枝数和第二级分枝数均减少、颖花数减少、籽粒数减少，最明显的特征是穗码变稀、单穗产量显著降低（$P<0.05$），但籽粒变大，千粒重极显著增大（$P<0.01$）。

*lp1* 的突变性状由位于 2 号染色体的细胞核隐性单基因 *LP1*（*Seita.2G369500*）控制，该基因属于 WRKY 转录因子家族的I亚族；*LP1* 基因的第五个内含子末端的碱基发生 G→A 的突变，导致 *lp1* 的 *LP1* 基因形成 3 个不同于野生型的可变剪切；但 3 个转录本均属于翻译提前终止型突变，造成 LP1 蛋白 C-末端第二个 WRKY 结构域的 C2H2 锌指结构遭到破坏，使 LP1 蛋白功能丧失。

*LP1* 基因分别在拔节期的茎节、孕穗期的穗、灌浆期的籽粒中高表达，这种表达模式与其调控株高、穗型和种子大小的功能一致；亚细胞定位表明，*LP1* 基因在细胞核中表达，符合转录因子基因的表达特征。

谷子的穗发育是受多个基因控制的复杂过程。本研究对谷子穗发育异常突变体 *lp1* 进行表型分析、目标基因图位克隆及功能分析，发现 WRKY 家族转录因子 *LP1* 突变会导致谷子穗发育异常，表现出明显的稀码特征。这些研究结果对揭示转录因子的生物学功能、解析谷子及禾本科植物穗发育的分子机制、指导谷子及其它作物高产育种等具有重要意义。

**关键词：谷子；稀码；转录因子；WRKY**

# 基因型和环境对烟草代谢产物的影响研究

金静静，刘萍萍，卢鹏，李泽锋，张剑锋，许亚龙，周会娜，曹培健*

国家烟草基因研究中心，中国烟草总公司郑州烟草研究院，郑州，450001

*E-mail：jinjingjing1218@126.com

**摘要：**为了了解基因型和环境对于烟草主要代谢产物的影响，我们研究了不同生态环境下、不同烟草品种代谢产物的变化规律。选取烟草三个主栽品种分别在中国三大产区种植，利用 GC-MS、LC-MS 等对烟草代谢产物进行鉴定，共鉴定出 700 多个代谢物，其中定性的约 500 个。通过 PCA 等分析发现，烟叶代谢产物具有明显的地区特征，同一地区不同品种具有相似的代谢特征。整体来说，对于烟草主要代谢产物，环境贡献大于生育期，基因型次之，彼此的互作贡献较小。而在环境因素中，又以温度的影响最为重要。因此对于未来的烟草育种，针对不同的代谢产物，应根据环境和基因型的不同效应，采取合理的调控措施。

**关键词：**烟草；基因型；环境；代谢组

# 紫花苜蓿逆境响应转录因子 NAC 基因的分离及其功能的研究

申玉华

赤峰学院生命科学学院/农业科学研究院，赤峰 024000

**摘要：** NAC 转录因子是高等植物特有的一类转录因子，参与高等植物的诸多生物学过程，如植物的生长发育、激素信号转导，同时还具有调控逆境胁迫应答等非生物胁迫和病原菌侵染等生物胁迫的功能，该特点决定 NAC 转录因子在植物育种实践中存在较高的应用潜力，即有可能通过改变一个 NAC 转录因子基因的表达而起到改良多个生物学性状的结果。本研究从紫花苜蓿中克隆了两个 NAC 家族基因 MsNAC2 和 MsNAC3，并对其功能及调控机理进行了研究。本研究将为进一步了解紫花苜蓿 NAC 家族基因表达调控机理和苜蓿耐逆分子机制提供重要参考和依据，同时为紫花苜蓿抗逆新种质的培育提供理论基础和技术支持。

**关键词：紫花苜蓿；NAC 转录因子；抗逆；调控机理**

# Next generation sequencing of DNA pools

Sheng Wang (王晟)

College of Life Information Science and Instrument Engineering, Hangzhou Dianzi University,

Hangzhou 310018

E-mail: bsmagic@hdu.edu.cn

**Abstract:** Large scale sequencing is still economically challenging despite falling costs of next generation sequencing. Pool-seq provides a cost-effective alternative to sequencing individuals separately. However, it creates new problems and challenges for accurate variant call and allele frequency estimation. We review several robust and reliable strategies and methods for pooled NGS.

**Keywords:** Targeted sequencing; Target enrichment HaloPlex; Single nucleotide variant; Next generation sequencing; Pooled sequencing

## Introduction

High-throughput DNA sequencing methods that allow genomic sequencing and analysis with unprecedented speed and detail. The massive capacity of next-generation sequencers can be harnessed for sequencing specific genomic regions in hundreds to thousands of individuals. Although advances in the field have led to substantial reductions in the cost of genome sequencing, it is still an issue to control the costs associated with different methods and strategies for sequencing for almost all genomic research projects, because they influence the scope and scale. Next Generation Sequencing (NGS) of pools of individuals (Pool-seq) provides a cost-effective alternative to sequencing individuals separately. Pool-seq is an alternative cost- and time-effective option in which DNA from several individuals is pooled for sequencing (Anand et al., 2016). But pooling of DNA creates new problems and challenges for accurate variant call and allele frequency (AF) estimation.

## NGS of pooled samples

Next-generation sequencing (NGS) has been demonstrated to be a powerful approach to overcome the wide clinical and genetic heterogeneity of disorders. Clinical implementation of sequencing has focused on customized sequencing of actionable genes, exons, or regions. A targeted-capture NGS strategy is clinically pragmatic, because it is scalable, is economical, and allows for deeper sequencing coverage compared to whole genome or whole exome approaches. Thus, many laboratories are using or considering

custom capture gene panels for diverse applications, including discovery, validation testing, or clinical-grade assay development. There are four popular library-building methods: SureSelect Custom Target Enrichment (Agilent Technologies), Haloplex (Agilent Technologies), SeqCap (Roche Diagnostics, Basel, Switzerland) and Nextera (Illumina, San Diego, CA), could be used for targeted sequencing for a targeted panel of couples of genes related to target disorders. A study have compared these four custom-targeted DNA capture methods with respect to uniform sequencing and variant calling, and demonstrates the potential biases of capture strategies due to differences in experimental procedures and probe design that may affect performance, including alignment and uniformity (Samorodnitsky et al., 2015). HaloPlex technology is widely used since it is a novel, fast, and specific method suitable for targeted sequencing of relatively small regions in many samples (Berglund et al., 2013).

As a cost-effective alternative to sequencing individuals separately, NGS of pools of individuals is becoming a popular strategy for characterizing variation in population genomic research, with the availability of custom-tailored software tools (Schlötterer et al., 2014). The cost per patient of next generation sequencing for detection of rare mutations may be significantly reduced using pooled experiments (Evangelista et al., 2016). Pooled DNA sequencing is a fast and cost-effective strategy to detect rare variants associated with complex phenotypes in large cohorts. This pooling technique consists in analyzing a mixture of DNA from a group of individuals, and assigning the discovered causative mutations to a single patient (Žilinskas et al., 2014). Many research projects utilize sequencing of pools containing multiple samples for the detection of sequence variants as a cost saving measure (Anand et al., 2016; Calvo et al., 2010; Schlötterer et al., 2014).

One of the disadvantages in using pooled sequencing data for variant detection is that, the variants found are not easy to be traced back to the original individual samples. Sometime, low throughout Sanger sequencing technology must be additionally used to attribute variants to a specific patient in a pool with more than 2 samples. To solve the indecision between patients in an identified group, in some cases, using some controlling strategies that replicate samples in pools (allocating the patient in two different pools) could decrease and even eliminate the number of low throughout experiments (Žilinskas et al., 2014). Some techniques have been proposed for the planning of pooled experiments and for the optimal allocation of samples into pools. A freely available web-oriented application with intuitive graphical user interface and novel algorithms, OPENDoRM (optimization of pooled experiments in NGS for detection of rare mutations) has been written for the planning of pooled NGS experiments (Evangelista et al., 2016; Žilinskas et al., 2014). OPENDoRM implemented three distinct strategies: (i) without Replica; (ii) with

Replica and (iii) Hybrid. A study presented a 3D-pooled DNA sample pooling strategy in which multiplexed semi-nested PCR is combined with NGS library construction for the sequencing of amplicons derived from samples , enabling the researcher to obtain DNA sequences of 512 individuals while only sequencing 24 samples (Chi et al., 2014).

Furthermore, there is a risk of not detecting rare variants in pools with a large number of individuals. To resolve this issue, many programs have been developed for calling common and rare variants in analysis of pooled next-generation sequencing (NGS) data (Schlötterer et al., 2014), including CRISP (Bansal, 2010), LoFreq (Wilm et al., 2012), VarScan (Koboldt et al., 2012), SNAPE (Raineri et al., 2012) , SNVer (Wei et al., 2011), GATK (DePristo et al., 2011), etc. SNAPE uses a Bayesian approach for minor allele frequency (MAF) computation and SNP calling in pools with good power while retaining a low false discovery rate (FDR). SNVer formulates variant calling as a hypothesis testing problem and employ a binomial–binomial model to test the significance of observed allele frequency against sequencing error. CRISP is a program designed to comprehensive read analysis for identification of SNPs from pooled sequencing. It applies a sophisticated set of techniques to distinguish between false variants coming from sequencing errors compared to those from real variant alleles and has been tested on a number of pooled targeted capture datasets generated using the Illumina sequencing platform. LoFreq is a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. It models sequencing run-specific error rates to accurately call variants occurring in <0.05% of a population. VarScan is a platform-independent mutation caller for targeted sequencing data. It employs a robust heuristic/statistic approach to call variants that meet desired thresholds for read depth, base quality, variant allele frequency, and statistical significance. GATK as a popular variant calling software, introduced a framework for variation discovery and genotyping using next-generation DNA sequencing data. A study has evaluated sensitivity and specificity for five programs (GATK, CRISP, LoFreq, VarScan, and SNVer) with regard to their ability to detect variants in synthetically pooled Illumina sequencing data to detect known true variants (Huang et al., 2015). GATK, CRISP, and LoFreq all gave balanced accuracy of 80 % or greater for datasets with varying per-sample depth of coverage and numbers of samples per pool. VarScan and SNVer generally had balanced accuracy lower than 80 %. CRISP and LoFreq required up to four times less computational time and up to ten times less physical memory than GATK did, and without filtering, gave results with the highest sensitivity. VarScan and SNVer had generally lower false positive rates, but also significantly lower sensitivity than the other three programs.

For pooled DNA sequencing experiments, the number of short reads produced during sequencing and belonging to each patient might be very different, due to the contribution of DNA from each patient in a pool is in general difference. And in a pool with the diploid individual number n, each single variant would only be represented in approximately 1/2n of the pooled sample reads. Since a certain percentage of reads is affected by base call and mapping errors, some rare variants of a specific patient could have representation rates lower than the sequencing error rate, and might be in reads that are unusable and therefore discarded from the analysis. In the worst case, it becomes increasingly likely that a singleton variant will not be sequenced at all. A rule of thumb to deal with pooled sequencing experiments is that the number of short reads covering each position of interest should be at least thirty multiplied by the number of samples in the pool (Žilinskas et al., 2014).

**References**


Anand, S., Mangano, E., Barizzone, N., Bordoni, R., Sorosina, M., Clarelli, F., Corrado, L., Martinelli Boneschi, F., D'Alfonso, S., and De Bellis, G. (2016). Next Generation Sequencing of Pooled Samples : Guideline for Variants' Filtering. Sci. Rep. *6*.

Bansal, V. (2010). A statistical method for the detection of variants from next-generation resequencing of DNA pools. Bioinforma. Oxf. Engl. *26*, i318-324.

Berglund, E.C., Lindqvist, C.M., Hayat, S., Övernäs, E., Henriksson, N., Nordlund, J., Wahlberg, P., Forestier, E., Lönnerholm, G., and Syvänen, A.-C. (2013). Accurate detection of subclonal single nucleotide variants in whole genome amplified and pooled cancer samples using HaloPlex target enrichment. BMC Genomics *14*.

Calvo, S.E., Tucker, E.J., Compton, A.G., Kirby, D.M., Crawford, G., Burtt, N.P., Rivas, M.A., Guiducci, C., Bruno, D.L., Goldberger, O.A., et al. (2010). High-throughput, pooled sequencing identifies mutations in NUBPL and FOXRED1 in human complex I deficiency. Nat. Genet. *42*, 851–858.

Chi, X., Zhang, Y., Xue, Z., Feng, L., Liu, H., Wang, F., and Qi, X. (2014). Discovery of rare mutations in extensively pooled DNA samples using multiple target enrichment. Plant Biotechnol. J. *12*, 709–717.

DePristo, M.A., Banks, E., Poplin, R.E., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. *43*, 491–498.

Evangelista, D., Zuccaro, A., Lančinskas, A., Žilinskas, J., and Guarracino, M.R. (2016). A web-oriented software for the optimization of pooled experiments in NGS for detection of rare mutations. BMC Res. Notes *9*.

Huang, H.W., Mullikin, J.C., and Hansen, N.F. (2015). Evaluation of variant detection software for pooled next-generation sequence data. BMC Bioinformatics *16*, 235.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. *22*, 568–576.

Raineri, E., Ferretti, L., Esteve-Codina, A., Nevado, B., Heath, S., and Pérez-Enciso, M. (2012). SNP calling by sequencing pooled samples . BMC Bioinformatics *13*, 239.

Samorodnitsky, E., Datta, J., Jewell, B.M., Hagopian, R., Miya, J., Wing, M.R., Damodaran, S., Lippus, J.M., Reeser, J.W., Bhatt, D., et al. (2015). Comparison of Custom Capture for Targeted Next-Generation DNA Sequencing. J. Mol. Diagn. JMD *17*, 64–75.

Schlötterer, C., Tobler, R., Kofler, R., and Nolte, V. (2014). Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. Nat. Rev. Genet. *15*, 749–763.

Wei, Z., Wang, W., Hu, P., Lyon, G.J., and Hakonarson, H. (2011). SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. Nucleic Acids Res. *39*, e132.

Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H., Khor, C.C., Petric, R., Hibberd, M.L., and Nagarajan, N. (2012). LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. *40*, 11189–11201.

Žilinskas, J., Lančinskas, A., and Guarracino, M.R. (2014). Application of Multi-Objective Optimization to Pooled Experiments of Next Generation Sequencing for Detection of Rare Mutations. PLoS ONE *9*.

# Transcriptome sequencing analysis of sunflower responsive to *Verticillium dahliae* infection and mining of resistance-related Genes

郭树春[1,2]，李前忠[1]，张艳芳[2]

[1]内蒙古大学物理科学与技术学院，呼和浩特，010021

[2]内蒙古农牧业科学院，呼和浩特，010031

**Abstract：** Sunflower (*Helianthus annuus L.*), an important oil crop and ornamental plant, is also considered as a potential biodiesel source. However, Sunflower Verticillium Wilt (SVW) is threatening the industry. In order to accelerate breeding progress, it is necessary to further study the molecular mechanisms of SVW, to find important genes and biological pathways related to the resistance. The research on molecular mechanism of resistance to SVW develops slowly for a lack of disease resistance materials and genetic information.

In this study, physiological and biochemical indexes were analysized in S18 and P77 inoculated by *Verticillium dahliae*. Furthermore, 16 samples that were infected by the pathogen in different time were sequenced by RNA-Seq. And a great deal of transcriptome information about sunflower infected with *V. dahliae* was obtained. And then, the study on interaction mechanism of Sunflower and SVW was made by bioinformatics methods. Based on the RNA-seq data, a lot of differentially expressed genes (DEGs) were detected by examining from different pairwise transcriptome comparisons. Disease resistance related DEGs were obtained by GO (Gene Ontology) enrichment analysis. Based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and correlation network analysis, the resistance mechanism of sunflower was explored. In general, based on the transcriptome profiling results, a detailed discussion on the transcripts about the crosstalk between sunflower and *V. dahliae* was provided, which was beneficial to disease-resistance breeding for sunflower.

**Keywords:** *Helianthus annuus L.*; RNA-Seq; Tanscriptome; Verticillium Wilt; Differentially expressed genes (DEGs); Resistance-related genes; qPCR validatio

# 蓖麻研究工作进展

黄凤兰[1,2,3]

[1]内蒙古民族大学生命科学学院，通辽 028043；

[2]内蒙古自治区高校蓖麻产业工程技术研究中心，通辽 028043；

[3]内蒙古自治区蓖麻育种研究重点实验室，通辽 028043

**摘要：**报告主要包括两个方面内容：一是平台简介，包括平台成立的背景、科研成果、科研团队、研究方向及研究内容。二是平台在蓖麻分子育种方面的研究进展。

# TCMID 2.0: a comprehensive resource for TCM

Lin Huang[1,#],Yan Shi[2,#], Yiran Yu[1], Duoli Xie[1], Huanlong Liu[2], Tieliu Shi[2,*], Chengping Wen[1,*]

[1]TCM Clinical Basis Institute, Zhejiang Chinese Medicine University, 548 Binwen Road, Hangzhou, Zhejiang, 310000, China; [2] Center for Bioinformatics and Computational Biology, and The Institute of Biomedical Sciences, College of Life Science, East China Normal University, Shanghai 200241, China.

* E-mail: wengcp@163.com

**Abstract:** As a traditional medical intervention in Asia and a complementary and alternative medicine in western countries, Traditional Chinese Medicine (TCM) is capturing worldwide attention in life science field. Traditional Chinese Medicine Integrated Database (TCMID), which was originally launched in 2013, was a comprehensive database aiming at TCM's modernization and standardization. It has been highly recognized among pharmacologists and scholars in TCM researches. The latest release, TCMID 2.0 (http://www.megabionet.org/tcmid/), replenished the preceding database with 18,203 herbal ingredients, 15 prescriptions, 82 related targets, 1,354 drugs, 1,572 diseases and numerous new connections between them. Considering that chemical changes might take place in decocting process of prescriptions, which may result in new ingredients, new data containing the prescription ingredients was collected in current version. In addition, 778 herbal mass spectrometry (MS) spectra related to 170 herbs were appended to show the variation of herbal quality in different origin and distinguish genuine medicinal materials from common ones while 3,895 MS spectra of 729 ingredients were added as the supplementary materials of component identification. With the significant increase of data, TCMID 2.0 will further facilitate TCM’s modernization and enhance the exploration of underlying biological processes that are response to the diverse pharmacologic actions of TCM.

**Keyword:** Traditional Chinese Medicine (TCM); database; mass spectrometry; network

# 磷在农田土壤中的迁移转化规律及其对农业环境的影响

张艾明

赤峰学院资源与环境科学学院/农业科学研究院，赤峰 024000

**摘要：**磷素作为植物体能量转化（ATP）、细胞结构（磷脂）、新陈代谢和信号传导的中心物质，是植物生长发育所必需的一种大量营养元素。磷素在自然界中主要以 5 价的形态存在，包括有机磷和无机磷两大组分，土壤中的无机磷必须经过解吸或溶解的过程，而有机磷必须经过一系列的矿化过程释放出无机磷才可被植物吸收利用。然而，在目前土壤生产管理中磷肥的当季利用率一般只有 10%-25%，施入土壤中未被植物吸收利用的磷肥部分被土壤吸附固定，而大量磷肥会随地表径流作用流入江河造成磷素的面源污染。因此，了解磷在农田土壤中的迁移转化规律对于磷肥资源的高效利用和保护环境都具有十分重要的意义。基于此，本研究主要分析了磷在农田土壤中的迁移转化规律以及土壤磷肥的施入对农业环境所造成的面源污染情况，包括磷肥的施用现状分析、土壤磷在土壤中的组分及其转化利用、磷对农业环境的影响等。旨在了解磷素在土壤中的形态分布及其迁移转化规律，提高农田土壤磷的利用率和降低磷在土壤中的累积和流失，从而减少磷对农业环境的污染，为农业的健康可持续发展提供指导。

**关键词：**磷；迁移转化；农业环境；面源污染

# Genome-wide identification and characterization of putative lncRNAs in the diamondback moth, *Plutella xylostella* (L.)

Yue Wang[1,2,3,4], Tingting Xu[1,2,3,4], Weiyi He[1,2,3,4], Xiujing Shen[1,2,3,4], Qian Zhao[1,2,3,4], Jianlin Bai[1,2,3,4], Minsheng You[1,2,3,4*]

[1]State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, Fujian Agriculture and Forestry University, Fuzhou 350002, China.

[2]Institute of Applied Ecology, Fujian Agriculture and Forestry University, Fuzhou 350002, China.

[3]Fujian-Taiwan Joint Centre for Ecological Control of Crop Pests, Fujian Agriculture and Forestry University, Fuzhou 350002, China.

[4]Key Laboratory of Integrated Pest Management for Fujian-Taiwan Crops, Ministry of Agriculture, Fuzhou 350002, China.

*E-mail: msyou@iae.fjau.edu.cn; Phone: 0086-591-8379-3035; Fax: 0086-591-8376-8847

**Abstract:** Long non-coding RNAs (lncRNAs) are of particular interest because of their contributions to many biological processes. Here, we present the genome-wide identification and characterization of putative lncRNAs in a global insect pest, *Plutella xylostella*. A total of 8,096 lncRNAs were identified and classified into three groups. The average length of exons in lncRNAs was longer than that in coding genes and the GC content was lower than that in mRNAs. Most lncRNAs were flanked by canonical splice sites, similar to mRNAs. Expression profiling identified 114 differentially expressed lncRNAs during the DBM development and found that majority were temporally specific. While the biological functions of lncRNAs remain uncharacterized, many are microRNA precursors or competing endogenous RNAs involved in micro-RNA regulatory pathways. This work provides a valuable resource for further studies on molecular bases for development of DBM and lay the foundation for discovery of lncRNA functions in *P. xylostella*.

**Keywords:** long-noncoding RNAs, DBM, RNA-seq, expression profiling

# 植物表观遗传修饰的多层次调控研究

王晶晶，陈铭*

浙江大学生命科学学院，杭州，310058

*E-mail：mchen@zju.edu.cn

**摘要:** 表观遗传修饰在基因表达调控中起着重要的作用，主要的机制包括组蛋白修饰、组蛋白变异、DNA 甲基化和非编码 RNA。本研究围绕表观遗传修饰的多层次调控这一研究方向，基因高通量测序数据和芯片数据，利用生物信息学方法开展全面深入的探索性研究。主要内容及结果概括为以下几点：

1、我们探索了基因重复发生后表观修饰模式的进化情况。整合多种组蛋白修饰图谱，我们检测了重复基因对的表观遗传分化、序列分化、表达分化和功能分化的四者之间的关系。结果显示重复基因对有相似的表观修饰模式；表观遗传分化与序列分化具有一致性；表观遗传分化与表达分化成正相关；功能分化对于表观遗传分化有一定的影响。

2、基于现有拟南芥在七类营养胁迫下的高通量测序数据，我们定义了 721 个 lncRNA 基因（874 个转录本），并分析了这些 lncRNA 序列特征、表观修饰特征和表达模式。通过构建包含 lncRNA 的 ceRNA 网络和共表达网络，预测 lncRNA 的功能。结合以上两个网络和差异表达模式，找出了关键 lncRNA，构建营养胁迫-关键 lncRNA 网络，系统分析关键 lncRNA 在多重营养胁迫下的调控作用。

3、我们整合了多组学的实验数据，包括 Hi-C、BS-seq、ChIP-chip/-seq 和 RNA-seq。利用 Hi-C 测序数据定义了拟南芥染色质长距离互作位点；结合组蛋白修饰谱，将染色质长距离互作位点分为七组；结合表达谱，将七组染色质长距离互作位点分为三个大类。结果显示互作的两个位点有相似的表观修饰模式和转录因子结合谱；七组染色质长距离互作位点的表观修饰模式不同并且转录因子结合的情况也不同；三大类的染色质长距离互作位点包含基因的表达模式不同：共同促进表达、共同抑制表达、无共调控表达。

通过以上研究，我们探索了不同机制的表观遗传的多层面的调控作用。希望能拓展目前人们对植物表观遗传调控的认识，为后续多种机制表观遗传协同作用调控的研究打下基础。

**关键词**：表观遗传调控，组蛋白修饰，组蛋白变异，DNA 甲基化，非编码 RNA

**致谢:** 相关研究受国家自然科学基金项目资助 (No. 31371328，31571366，31611130033)。

**参考文献:**

Wang J, Meng X, Chen H, Yuan C, Li X, Zhou Y, Chen M (2016) Exploring the mechanisms of genome-wide long-range interactions: interpreting chromosome organization. Briefings in Functional Genomics 15: 385-395

Wang J, Meng X, Yuan C, Harrison AP, Chen M (2016) The roles of cross-talk epigenetic patterns in Arabidopsis thaliana. Briefings in Functional Genomics 15: 278-287

Wang J, Zhou Y, Li X, Meng X, Fan M, Chen H, Xue J, Chen M (2016) Genome-Wide Analysis of the Distinct Types of Chromatin Interactions in Arabidopsis thaliana. Plant and Cell Physiology 58: 57-70

Yuan CH, Wang JJ, Harrison AP, Meng XW, Chen DJ, Chen M (2015) Genome-wide view of natural antisense transcripts in Arabidopsis thaliana. DNA Research 22: 233-243

# 陆生植物 *TAL* 基因的水平转移起源及对水稻维管束发育和产量性状的作用

杨泽峰，周勇，梁国华*，徐辰武*

扬州大学农学院/植物功能基因组学教育部重点实验室/江苏省作物遗传生理重点实验室/粮食作物现代产业技术协同创新中心，扬州，225009

*E-mail：ghliang@yzu.edu.cn; qtls@yzu.edu.cn

**摘要：**维管束是高等植物的重要特征，其在植物体内交织连接构成维管系统，起着为植物体疏导水分、无机盐和有机物质的作用，并兼有支持植物体的作用。微管组织的出现，一方面使植物植株变得高大，另一方面使植物真正意义上摆脱了水生环境，能够在复杂恶劣的陆地上生存。维管系统的有效输导，使维管植物成为最繁茂的陆生植物。然而，有关植物维管组织起源和进化的分子机制目前尚不明确。

转醛醇酶（transaldolase）作为磷酸戊糖途径（PPP）非氧化阶段的关键酶，一般认为其在调节植物 PPP 对环境胁迫的应答中起重要作用。本研究对 TAL 类型的植物转醛醇酶基因进行了系统的生物信息学和进化分析，并在水稻中对该基因进行了功能鉴定。

同源性搜索发现 *TAL* 基因在已经测序的陆生植物（Land plants，Embryophyta）基因组中普遍存在，但在陆生植物祖先物种绿藻及其他真核生物生物中均没有同源基因。进一步对 *TAL* 基因进行了系统的进化分析，结果发现该基因的同源基因在细菌中广泛存在，并且陆生植物的 *TAL* 基因在系统进化树中位于放线菌（Actinobacteria）的分枝中，由此表明，*TAL* 基因最初起源于细菌，而陆生植物的 *TAL* 基因是经一次水平基因转移（Horizontal gene transfer，HGT）事件起源于细菌，供体为放线菌。进一步的研究还发现：*TAL* 基因自细菌转移至陆生植物的祖先物种之后，该基因经历了明显的正选择作用，可能促使该基因在植物中产生了不同于细菌同源基因的新的功能。

实时定量 qRT-PCR 分析发现水稻 *TAL* 基因在根、茎、叶、花和节等器官中均表达；而 GUS 染色分析表明水稻的 *TAL* 基因特异性地在维管组织中表达。分别构建了水稻 *TAL* 基因的 RNA 干涉和过表达转基因株系，并对该基因进行了初步的功能鉴定。RNA 干涉转基因植株表现出株高降低，叶片卷曲等表型特征，茎秆和叶片中的维管束数目减少、维管束的形态变小；水稻 *TAL* 基因的过量表达株系能够在缩短生育期的前提下，显著增加水稻千粒重、每穗粒数、单株籽粒产量和生物学产量以及群体籽粒产量。

本研究的结果表明：水平转移起源的陆生植物 *TAL* 基因在维管束发育中具有重要作用，其起源和进化对于植物适应陆地生态环境具有至关重要的作用；此外，该基因在农作物产量性状的基因工程育种中具有较大的应用潜力。

**关键词：**类转醛醇酶；陆生植物；水平基因转移；维管组织；产量

# Crucial Enzymes in the Hydroxylated Triacylglycerol-ricinoleate Biosynthesis Pathway of Castor Bean

Yujie Chen[1,2], Lili Liu[3,*], Xun Tian[1], Jianjun Di[1,2], Yalatu Su[1,2], Fenglan Huang[1], Yongsheng Chen[1,*]

[1] Inner Mongolia Molecular Breeding Laboratory of Industrial Engineering Research Center for Castor, College of Life Sciences, Inner Mongolian University for the Nationalities, Tongliao 028000, China;

[2] Institute of Bioinformatics and Systems Biology, Inner Mongolian University for the Nationalities, Tongliao 028000, China; 3 Department of Bioinformatics, College of Life Sciences, Zhejiang University, Hangzhou 310058, China

*E-mail: llliu@zju.edu.cn; chenyongsheng@imun.edu.cn

**Abstract:** Castor bean (*Ricinus communis* L.) is an important oilseed crop for the rich hydroxylated triacylglycerol (TAG)-ricinoleate which is a raw material with wide applications in industry. Hydroxylated TAG synthesis occurs through complicated pathways among multiple subcellular organelles. Some crucial enzymes have been identified in previous studies. After analyzing the available castor tissue-specific transcriptome sequencing data and comparing the classic pathways in other plants, a possible *de novo* biosynthesis pathway for the hydroxylated TAG has been revealed. In this study, some other crucial enzymes were ascertained and their expression levels were improved and pinpointed into the pathways in castor. Several key enzymes were analyzed in terms of structure, biofunction prediction and similarity of expression pattern mechanisms, aiming to give an insight on the better understandings of the molecular knowledge for this oil-rich plant and the crucial enzyme performances in the hydroxylated triacylglycerol-ricinoleate biosynthesis pathways.

**Keywords:** Biosynthesis pathways, castor bean, hydroxylated TAG-ricinoleate, Transcriptome sequencing data.

# 蓖麻醇酸酯生物合成代谢途径及关键酶研究进展

狄建军[1,2]，黄凤兰[1,2]，陈永胜[1,2]，张树军[1,2]，魏永春[1]

[1]内蒙古民族大学生命科学学院，通辽 028043

[2]内蒙古自治区高校蓖麻产业工程技术研究中心，通辽 028043

**摘要：**蓖麻(*Ricinus communis* L．) 是起源于非洲的热带多年生灌木，属大戟科的一种油料植物，在世界上热带及亚热带区域种植。现在蓖麻基因组草图绘制完成，在大戟科成员中尚属首例。随着近年来分子生物学和基因工程技术的发展，蓖麻醇酸酯合成代谢途径基本确定，研究主要对蓖麻醇酸酯生物合成代谢途径及关键酶——脂酰－ACP 硫酯酶、油酰－12 羟化酶、二酰甘油酰基转移酶、磷脂：二酰甘油酰基转移酶、磷脂酰胆碱二酰甘油胆碱磷酸转移酶等几个关键酶进行综述，对蓖麻醇酸酯代谢遗传工程应用研究有较大的意义。

# 通蓖 5 号蓖麻籽粒干物质及粗脂肪积累规律研究

狄建军[1, 3, 4]，张树军[1, 3, 4]，陈宇杰[1,3,4]，朱小慧[1]，王云[2,3]

[1]内蒙古民族大学生命科学学院，通辽 028043；

[2]内蒙古民族大学农学院，通辽 028043；

[3]内蒙古自治区高校蓖麻产业工程技术研究中心，通辽 028043；

[4]内蒙古自治区蓖麻育种研究重点实验室，通辽 028043

**摘要：**以大穗型品种通蓖 5 号蓖麻为材料，研究了蓖麻籽粒干物质及粗脂肪积累的规律，结果表明，授粉后随着时间的推移，蓖麻籽粒干重积累量动态呈现“慢—快—慢”的生长方式，即符合“S”型曲线生长方式,在 35 DAP 时百粒干重为 26.86g，接近最大值 28.80g.籽粒每克干重含油量在 50 DAP 时达到最大值，籽粒含油率为 52.80%.

# 褪黑素介导绒山羊皮肤比较转录组研究

丽春[1,*]，付少印[2]，李金泉[3,*]，张文广[3,*]

[1]内蒙古民族大学动物科学技术学院，通辽 028043；

[2]内蒙古农牧业科学院，呼和浩特 010031；

[3]内蒙古农业大学动物科学技术学院，呼和浩特 010018

*E-mail：lijinquan_nd@126.com，atcgnmbi@aliyun.com，lichun1985@126.com

**摘要：**为探索内蒙古绒山羊绒毛周期性变化的重要分子事件，阐明褪黑素影响绒山羊绒毛周期性变化的分子机理，选择 6 只罕山白绒山羊，分对照组和埋植组（耳后皮下持续埋植褪黑素），0-12 个月（0 月为实验开始时间，即上一年的 12 月）间每月连续采集其皮肤样品，共获得 78 个样品进行转录组测序。对转录组 cleandata 数据使用 TCC 流程比对到山羊参考基因组 Caprahircus (assembly ARS1) ，序列成功比对率为 78.38+/-1.69%。

0-12 个月的对照组样品根据 FPKM 值进行不同月份基因的多重假设检验发现，其 7 月对比 8 月差异表达基因比其他月份具有显著性差异现象，对此进行富集分析发现， 差异表达基因主要富集在 Jak-STAT signaling pathway 、 Complement and coagulation cascades、 Measles、 Fatty acid metabolism 等信号通路；然而埋植组此时变化不明显。

0-12 个月对照组和埋植组差异基因数目变化关系分析发现，4-9 月份差异基因数目显著增加，其它月份变化不明显。根据两组各月份间差异表达分析推测，GSMTT 为褪黑激素潜在靶基因，褪黑素增加其表达量，完全改变了与其它基因相互作用的关系。

以转录组 SNP 分析为切入点，通过计算位点的 Fst 值，进一步分析褪黑素介导的等位基因特异表达情况。共检测到 SNP 位点 11810021，Fst 值>0.6 的位点为 398633。取 25 号染色体 Fst 值>0.6 的位点，25821 个位点映射到了 128 个基因。通过 DAVID（https://david.ncifcrf.gov/）在线分析工具对这 128 个基因进行功能富集分析发现，这些基因富集到三个通路：NF-kappa B signaling pathway、 Rap1 signaling pathway、 Retrograde endocannabinoid signaling。

**结论：**7 月、8 月是罕山白绒山羊皮肤毛囊发育的重要转折点；4 月、9 月是褪黑素影响皮肤毛囊发育的关键时期；GSMTT 基因是皮肤中褪黑素作用的一个潜在靶点。

**关键词：褪黑素，绒山羊皮肤组织，转录组**

**基金资助：**内蒙古自然科学基金博士项目(No.2014BS0313)； 国家自然科学基金项目(No.31402052)

# Dissecting the Underlying Pharmaceutical Mechanism of Chinese Traditional Medicine Yun-Pi-Yi-Shen-Tong-Du-Tang Acting on Ankylosing Spondylitis through Systems Biology Approaches

Duoli Xie [#], Lin Huang [#], Yiran Yua, Jiawei Gao, Haichang Li, Chengping Wen [*]

TCM Clinical Basis Institute, Zhejiang Chinese Medicine University, 548 Binwen Road, Hangzhou, Zhejiang, 310000, China

[#] These authors contributed equally to this work.

∗ Correspondence should be addressed to C.W. E-mail: wengcp@163.com; Tel: +86-571-86613644.

**Abstract:**

Traditional Chinese Medicine (TCM) has been served as complementary medicine for Ankylosing Spondylitis (AS) treatment for a long time. Yun-Pi-Yi-Shen-Tong-Du-Tang (Y-Y-T)[1] is one of the effective formulas and has been widely used in clinical for AS treatment. In this study, a clinical practice proved the efficiency of Y-Y-T in AS treatment and then we tried to decipher the underlying molecular mechanism of Y-Y-T in the therapy. Herbal ingredients and targeting proteins were collected from available databases TCMID and STITCH. PPI networks were then constructed to further infer the relationship among Y-Y-T, drugs used for treating AS, differentially expressed genes of AS patients and AS disease proteins. We found that TLR signaling pathway and T cell receptor signaling pathway may involve in the biological processes of AS progression and contribute to the curative effect, proteins such as JAK2, STAT3, HSP90AA1, TNF and PTEN are the key targets in the therapy. Our systemic investigation to infer therapeutic mechanisms of Y-Y-T for AS treatment provides a new insight in the understanding of TCM functions.

**Keywords:** Ankylosing Spondylitis; Traditional Chinese Medicine; formula; pharmacology network

## Introduction

Ankylosing Spondylitis (AS) is a chronic rheumatic disease with 0.2-0.5% prevalence worldwide [1]. It mainly affects the axial skeleton and is characterized by morning stiffness and sacroiliitis [2]. In the disease process, the onset of joint fusion, spinal deformity and disability will sequentially take place [3]. So far, the

1 Y-Y-T; Yun-Pi-Yi-Shen-Tong-Du-Tang

accurate etiology of AS is still unknown. It was reported that the etiopathogenesis of AS was related to Gram-negative bacteria, human leukocyte antigen B27 (HLA-27), pattern recognition receptors (PRRs) and inflammatory bowel disease (IBD) [4,5]. AS remains an incurable disease. Medication treatment for AS mainly includes the following three classes of drugs, non-steroid anti-inflammatory drugs (NSAIDs), disease-modifying anti-rheumatic drugs (DMARDs) and biologicals. These medical treatments can alleviate inflammatory reaction, relieve pain of sacroiliac joints and spine, slow down disease progression and decrease disease activity to some extent. But a prolonged therapy with these medications may also cause systematic side effects such as serious infections and gastrointestinal intolerance [6,7].

Traditional Chinese Medicine (TCM), as an alternative medicine, is widely used for AS treatment in clinical. For example, kunxian capsule, a Chinese patent medicine used for immunologic treatments, can effectively induce an anti-inflammatory effect and regulate immunity [8]. Besides, clinical observation found that Chinese herbs combined with etanercept could relieve various symptoms and improve therapeutic efficacy in AS patients [9]. Another clinical case reported that Bushen-Qiangdu-Zhilv Decoction can help AS patients alleviate the inflammatory symptoms and improve quality of life [3].

Yun-Pi-Yi-Shen-Tong-Du-Tang (Y-Y-T) is a TCM decoction created by Professor Chengping Wen. It contains 11 medicinal herbs with recommended doses as follows: *Dioscoreae Nipponicae Rhizoma* (Chuan Shan Long, 20g), *Atractylodes Lancea* (Cang Zhu, 12g), *Rhizoma Smilacis Glabrae* (Tu Fu Ling, 30g), *Lonicera Japonica* (Jin Yin Hua, 15g), *Achyranthes Bidentata* (Niu Xi, 12g), *Myrrh* (Mo Yao, 10g), *Aconitum Carmichaeli* (Chuan Wu, 10g), *Radix Astragali* (Huang Qi, 15g), *Glycyrrhiza* (Gan Cao, 6g), *Leech* (Shui Zhi, 6g), *Coptis* (Huang Lian, 9g). Studies have shown that *Myrrh*, *Coptis* and *Lonicera Japonica* were remedies for inflammation related disorders [10-12]. *Aconitum Carmichaeli* is an analgesic and anti-rheumatic medicine which can effectively alleviate the symptoms of neuropathic pain and inflammatory [13] while *Dioscoreae Nipponicae Rhizoma* has been widely used to deal with arthroncus, arthrodynia and arthritis [14].

Network pharmacology has provided a new sight to understand the mechanisms of TCM and is widely used in omics and bioinformatics in recent years [15,16]. For instance, systemic pharmacology-based approach was applied to predict potential targets and molecular mechanism, for example, Bu Fei Jian Pi formula can act on the COPD [17] and to investigate pharmacological mechanism of how Wu Tou Tang take effect on Rheumatoid Arthritis [18]. AS is a disease with complex pathogenic factors. Y-Y-T showed its own worth in the intervention of disease progression. To clarify the underlying mechanism of Y-Y-T in AS treatment, we deeply analyzed the composition of the formula and constructed PPI networks to show

the interrelationship between formula targets and AS-related proteins and to predict potential targets of Y-Y-T. This systematic approach to uncover the mechanism of therapy on molecular lever facilitates our understanding of the intangible biological processes of this formula.

## Results

### Clinical performance of Y-Y-T in AS patients' treatment

We made a retrospective investigation to evaluate the effectiveness of the formula. In this study, patients who did not response to the western medicine therapy and continuously suffer from morning stiffness, persistent lower-back and multiple joints pain participated Y-Y-T add-on treatment for 3 times a day in 6 months. Inflammatory markers, acute phase reactants C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), Ankylosing Spondylitis Disease Activity Index ASDAS-CRP and ASDAS-ESR were measured before and after treatment. Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) was assessed and 9 patients were in active period of AS (BASDAI score of ≥4) at baseline. After the 6-month treatment, the data of BASDAI, ASDAS-CRP and ASDAS-ESR scores were significantly decreased (p-value as 0.0003, 0.003 and 0.002 respectively). Moreover, symptoms such as pain, fatigue and morning stiffness were apparently relieved. The result of clinical research showed that Y-Y-T formula could help AS patients in alleviating symptoms and enhancing physical quality.

### The herbs, ingredients and targets of the Y-Y-T

It is generally recognized that spleen, kidneys, and governor vessel, in the cognition of TCM, are closely related to the risk of AS progression. So, the creative principle of Y-Y-T is to strengthen spleen, nourish kidney and dredge governor vessel. Specifically, *Atractylodes Lancea, Radix Astragali* and *Glycyrrhiza* are used to strengthen the spleen, *Achyranthes Bidentata* is added to burst the function of kidneys while *Rhizoma Smilacis Glabrae* and *Aconitum Carmichaeli* are used to dredge governor vessel. The six herbs above play vital roles in the treatment and thus being called as Jun (Monarch) while the other 5 herbs are served as assistance. To further mine the underlying treatment mechanism of this formula, we collected herbal ingredients from TCMID resulting in 357 kinds of ingredients. To be detailed, *Aconitum Carmichaeli* contained 2 ingredients, gaconitine and songorine. *Myrrh* contained 2 ingredients, eugenol and cinnamaldehyde. *Dioscoreae Nipponicae Rhizoma* contained 5 ingredients, including dioscin, trillin, allantoin, etc. There were 68, 33, 106, 2, 70, 11, 35 and 30 ingredients in *Atractylodes Lancea*, *Rhizoma Smilacis Glabrae*, *Lonicera Japonica*, *Achyranthes Bidentata*, *Radix Astragali*, *Glycyrrhiza*, *Leech* and

*Coptis* respectively. Function of some ingredients has been studied. For instance, discin, one ingredients of *Dioscorea nipponica*, could repair the damaged synovium tissue by reducing Th1/Th2 [19], regulate the signaling pathway of the microRNA let7i/TLR4/MyD88 and reverse the inflammatory kidney injury [20]. Astilbin, a bioactive compound extracted from *Rhizoma Smilacis Glabrae*, could decrease antigen-specific autoantibodies by up-regulating regulatory T cells and down-regulating Th17 cells and it also can reduce the efficiency of antigen presenting cells by decreasing the expression of MHC class II [21].

Among the 357 ingredients, 350 of them were unique to individual herb while 7 were shared among herbs which indicated a cumulative effect (Figure 1). Mutual ingredients between 5 herbs formed 7 herb pairs. For instance, *Lonicera Japonica* and *Radix Astragali* shared 2 compounds γ-sitosterol and β-sitosterol, *Radix Astragali* and *Atractylodes Lancea* shared adeninenucleoside and uridine while chlorogenic acid was found in *Lonicera Japonica* and *Coptis*. Interestingly, most of the common ingredients can reduce inflammation through various approaches such as decreasing tumor necrosis factor (TNF) and interleukin 6 (IL-6) production, preventing leukocyte extravasation and antibacterial effect [22-25].

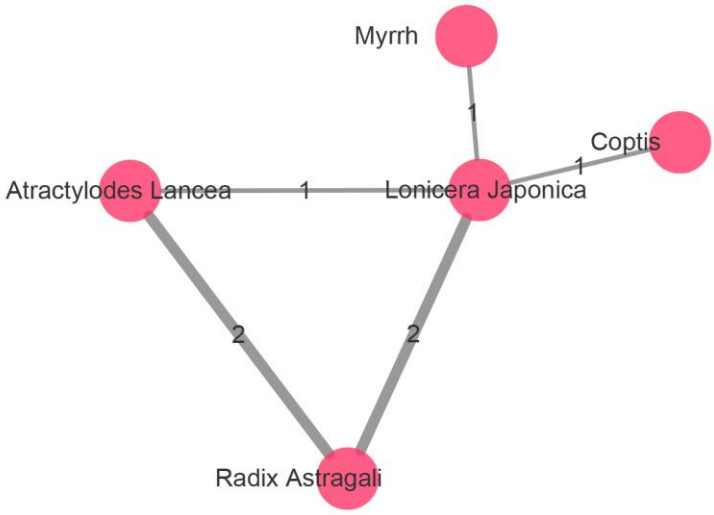


**Figure 1.** The relationship between herbs. Same ingredients shared between 5 herbs, the width of the lines represent the number of same ingredients shared between herbs.

We then extracted targets that were highly correlated with the ingredients from STITCH. Totally, we found that 81 ingredients had effect on 1100 proteins. 452 targets are shared among ingredients while other 648 targets are unique. Several small molecules such as resveratrol, trans-resveratrol and quercetin seem to have similar features since they possessed more than 70 identical targeting proteins (Figure 2).

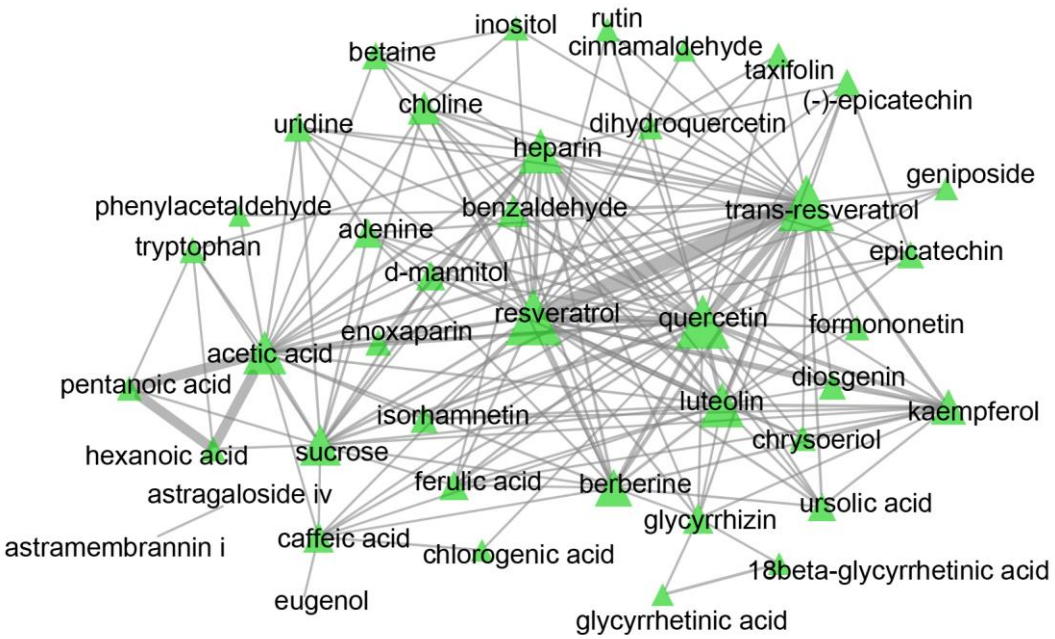


**Figure 2.** The relationship between ingredients. Sharing targets more than 3 between ingredients were showed up in the network. The width of the lines represents the number of same targets among ingredients.

**Functional analysis of Y-Y-T targets**

Gene Ontology (GO) analysis and KEGG pathway analysis were utilized to analyze the main therapeutic effect of the formula. Results of GO analysis showed that targets of Y-Y-T were mainly related to 5 highly parts, namely catalytic, transporter, receptor and molecular transducer activity as well as binding (Figure 3). Specifically, heparin binding ($p$ = 1.20E-11), peptide receptor activity, G-protein coupled ($p$ = 1.20E-09), neuropeptide hormone activity ($p$ = 3.10E-06) are closely associated with AS. For example, a neuropeptide of vasoactive intestinal polypeptide plays a significant role in neuroendocrine-immune-gastrointestinal systems [26] and magnesium-containing intramedullary nails had the ability to repair osteoporosis [27]. The enriched terms showed us the essential function of the formula in treating AS.

80 enriched KEGG pathways could be divided into six categories including Human Diseases, Organismal Systems, Environmental Information Processing, Metabolism, Genetic Information Processing and Cellular Processes. In Human Diseases group, Inflammatory bowel disease (IBD) ($P$ = 5.10E-03) and Rheumatoid arthritis (RA) ($P$ = 1.00E-02) are significantly associated with AS. The etiopathogenesis of IBD and AS is analogous because of the similar genetic and immunologic background. Moreover, RA and AS patients shared common bone metabolism biomarkers and symptoms [5,28]. Additionally, some pathways in immune system were involved in the onset and development of AS. For

example, in an inflammatory environment the activation of AMPK ($p$ = 5.00E-04) and TLR signaling pathway ($P$ = 7.30E-04) would lead to the pathogenesis of AS [29]. T cell receptor signaling pathway ($P$ = 9.10E-04) participated in hip joint ligament ossification [30]. Moreover, other pathways focus on endocrine, sensory and nervous systems and relate to signaling molecules interaction and transduction. In metabolism, the formula may involve in the metabolism process of amino acid, carbohydrate and lipid.

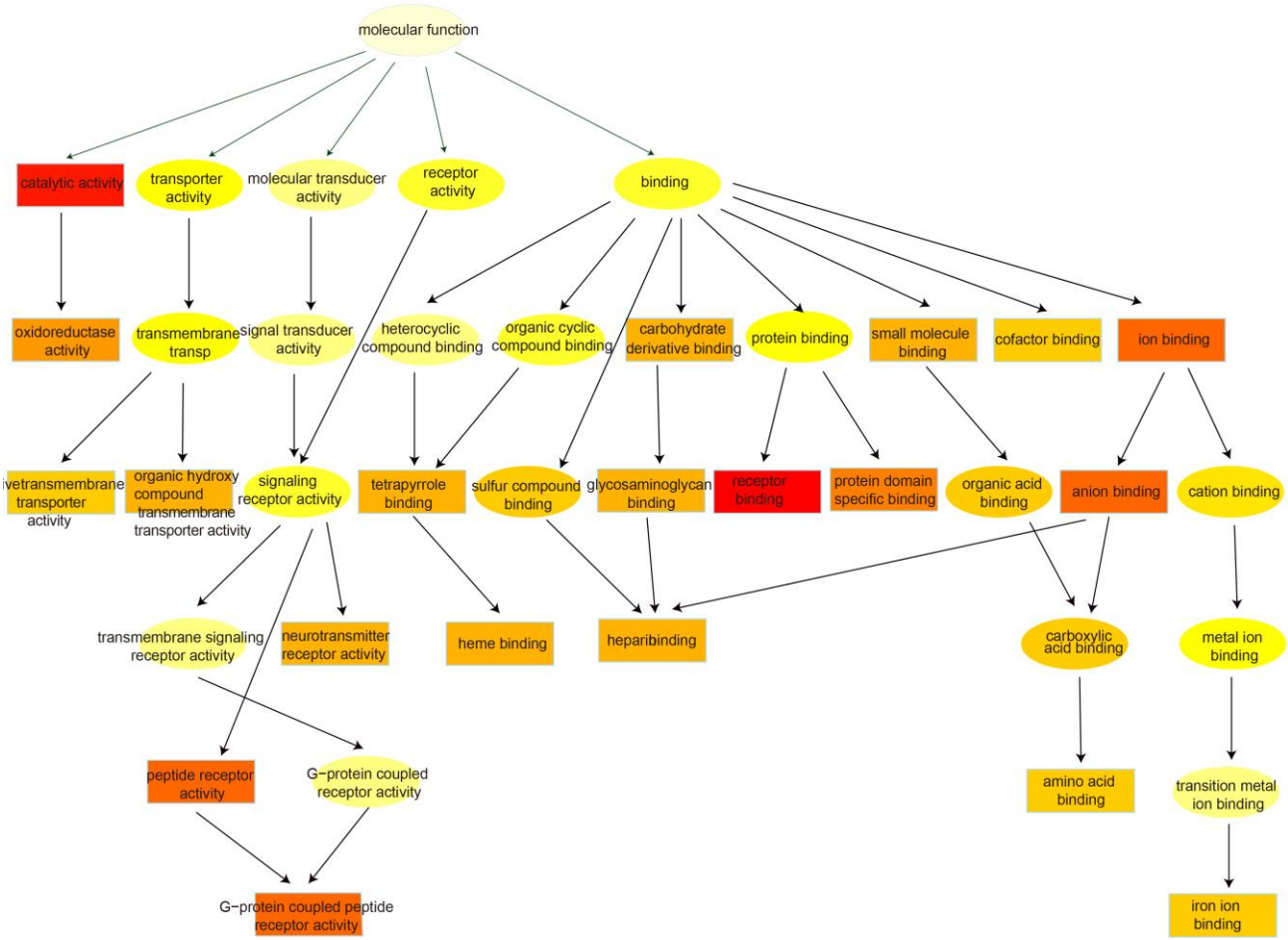


**Figure 3.** Classification of the formula's top 20 enriched Go term. The different colors represent different functional categories and the lower levels affiliation to upper levels.

**Comparison between Y-Y-T targets, other AS drug targets, AS disease proteins and differentially expressed genes of AS patients**

To further discover potential therapeutic mechanism of Y-Y-T formula for AS treatment, we made comparison between four datasets. The expression matrix sample of 72 AS patients were downloaded from GEO database. T-test with Benjamin-Hochberg correction for multiple comparisons was carried out to identify DEGs, resulting in 2122 DEGs. The other 3 datasets, 1100 targets of Y-Y-T formula, 88 targets of 19 FDA approved drugs related to AS and 115 AS disease genes, were retrieved from related public

databases. As showed in Figure 4. 19, 98 and 36 formula targets overlapped with disease proteins, proteins encoded by DEGs and drug targets, respectively. TNF, which is the only one target shared the 4 groups, is apparently a key protein in therapy.

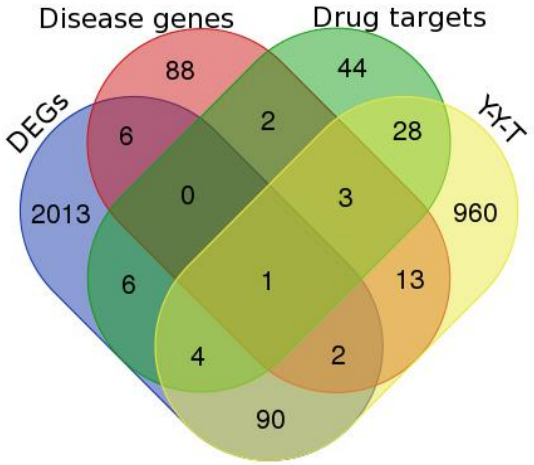


**Figure 4.** The overlaps between formula targets, disease genes, drug targets and DEGs. Blue: disease genes; Yellow: DEGs; pink: Y-Y-T formula targets; green: drug targets.

**Targets association networks between proteins encoded by disease genes and Formula targets**

To further explore the underlying mechanism of the formula, we applied PPI networks to uncover the functional relationship between formula targets and disease proteins. 811 formula targets and 77 proteins encoded by disease genes formed 3732 pairs of PPIs. As showed in Figure 5, two proteins overlapped between the two groups (janus kinase 2 and signal transducer and activator of transcription 3) were defined as essential targets. 17 nodes (13 formula targets, 2 essential targets and 2 AS disease genes), formed a highly-connected cluster. In this cluster, JAK2 and Epidermal growth factor receptor (EGFR) appear to act as hub due to their high degree of association. JAK2 and STAT3 are both regulatory factors of IL-23 pathway which is an important etiological factor for AS [31] while EGFR has been recognized as a monoclonal antibody targets for the treatment of AS [32] and involves in AS progression. Moreover, STAT3 is also an essential protein in differentiation and maintenance of Th17 cells. Furthermore, JAK2, signal transducer and activator of transcription 1 (STAT1), SHC-adaptor protein (SHC), protein tyrosine kinase 2 (PTK2) jointly take part in chemokine signal pathway which has been proven to affect immune-mediated inflammatory disease. However, heat shock protein 90 alpha family class A member 1 (HSP90AA1) is a formula targets directly or indirectly connected with 4 disease proteins. It was suggested that HSP90AA1-targeted agents can potently balance the inflammation-immune system through blocking inflammation, cytokine production, protein kinase activity and angiogenic signaling [33].

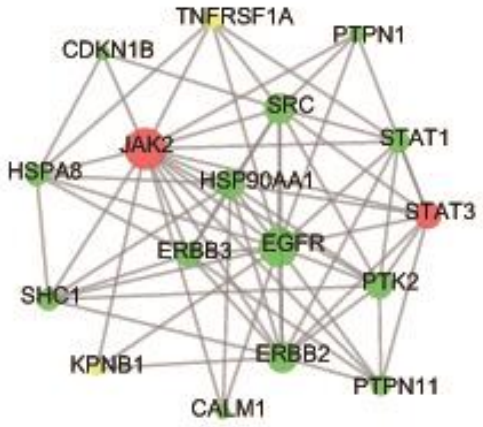


**Figure 5.** A cluster PPI network of the Y-Y-T formula and disease genes. Green: formula targets; yellow：disease genes; red: formula and disease genes.

**Targets association networks between DEGs and Formula targets**

869 Y-Y-T targets and 1266 DEGs formed 13461 pairs of PPIs and there are 90 targets in common. 79% Y-Y-T targets associated with 60% DEGs. For 90 Y-Y-T targets sharing with the DEGs, 41 of them were up-regulated DEGs while 49 were down-regulated DEGs. As showed in Figure 6, spectrin beta, non-erythrocytic 1 (SPTBN1), shared by 2 groups, was a hub connected with other 25 nodes including estrogen receptor 1 (ESR1), neurotrophic receptor tyrosine kinase 1 (NTRK1) and phosphatase and tensin homolog (PTEN), etc. ESR1 was related to low bone mineral density (BMD) in AS [34,35]. NTRK1 is involved in the pain mechanism [34] while PTEN could active PI3K/Akt signaling which is related to the bone metabolism of AS patients [36]. The formula may act on above mentioned proteins to intervene the progression of AS.

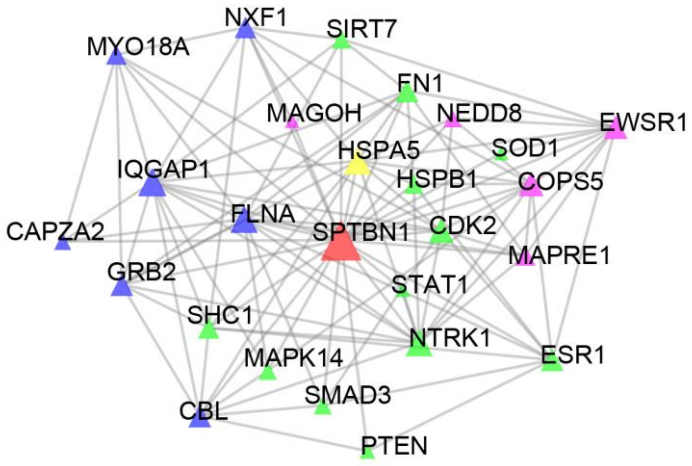


**Figure 6.** A cluster PPI network of the Y-Y-T formula and DEGs. Green: formula targets; red： formula targets and up-regulated DEGs; yellow: formula targets and down -regulated DEGs; blue: up-regulated DEGs; purple: down -regulated DEGs.

**Targets association networks between drug targets and Formula targets**

810 Y-Y-T targets interact with 68 drug targets and 34 drug targets can be found in formula targeting proteins, including TNF and prostaglandin G/H synthase 2 (PTGS2) in Etanercept, NFKB inhibitor alpha (NFKBIA) and inhibitor of nuclear factor kappa B kinase subunit beta (IKBKB) in Acetylsalicylic acid. These drugs are routine medication in the course of AS treatment, indicating an akin mechanism in formula. As presented in Figure 7, hub target TNF was connected with other 10 targets. TNF is a cytokine involved in systemic inflammation and anti-TNF drugs are widely utilized for rheumatism intervention in clinical [37]. Another drug target NFKBIA and its' promoter polymorphisms are associated with the development of AS [38]. In addition, formula target conserved helix-loop-helix ubiquitous kinase (CHUK) has been reported to effect as anti-TNF [39] and help control the inflammation. All these results expounded that a combination of AS drug targets and Y-Y-T formula targets can enhance the curative effect.

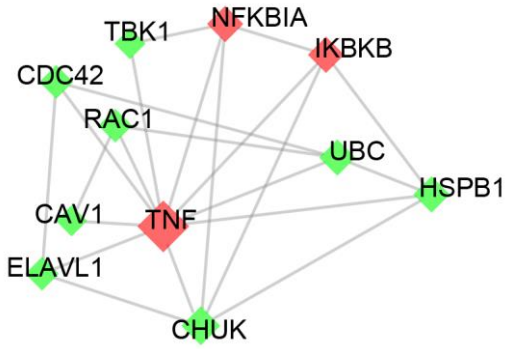


**Figure 7.** A cluster PPI network of the Y-Y-T formula and drug targets. Green: formula targets; red: formula targets and drug targets.

**Discussion**

Y-Y-T formula is a prescription created by Professor Chengping Wen. Clinical evidence demonstrates that Y-Y-T formula is beneficial for AS patients in relieving clinical symptoms and enhancing physical quality. In this study, we proved the efficacy of Y-Y-T on AS treatment and deciphered the potential underlying mechanism of this formula.

BASDAI is a self-administer instrument consisting of six horizontal visual analog scales and is widely used to estimate the disease progression and measure severity of fatigue, spinal, and peripheral joint pain, localized tenderness and morning stiffness [40]. ESR and CRP are biomarkers of inflammation and are widely utilized in various diseases for evaluation of disease activity and medicine efficacy [41], in addition, an elevated CRP is associated with increased damages of iconography both on X-rays and sacroiliac MRI [42,43]. The retrospective research result indicated that the formula Y-Y-T can reduce the symptoms of

morning stiffness, fatigue, pain and decrease clinical indexes of disease activity, proving the effectiveness.

In this study, herbs, compounds and targets of Y-Y-T formula were jointly exploited in layers of in-depth analysis, which gave us a better understanding of the potential therapeutic mechanisms of Y-Y-T. We found that there are a few common compounds among the 11 herbs. Some compounds, like resveratrol, trans-resveratrol, quercetin targeting more than 70 proteins seems to be critical compounds in the therapy. For example, resveratrol can restrain growth of aklebsiella pneumonia [44] which is an inflammation trigger of AS while quercetin has anti-inflammatory and analgesic effects [45]. Meanwhile, some targets were shared among the compounds, such as IL-6, nitric oxide synthase 3 (NOS3), apolipoprotein B (APOB), TNF, PTGS2, matrix metallopeptidase 9 (MMP9), caspase 3 (CASP3) and cytochrome P450 family 1 subfamily B member 1 (CYP1B1), indicating an accumulative effect. IL-6, TNF, PTGS2 are classic pro-inflammatory cytokines secreted by a variety of immune cells and are either associated with disease activity or associated with other inflammatory markers [46,47]. However, berberine and astibin could inhibit the expression and secretion of IL-6, TNF and PTGS2 [48-50].

Functional analysis illustrated that formula was significantly enriched on GO terms of molecular transducer activity, receptor activity as well as binding and pathways related to immune, inflammatory and metabolism, like metabolic pathways, TLR signaling pathway, AMPK signaling pathway, T cell receptor signaling pathway, etc. Most of these pathways were involved in the development of AS. Others, such as Complement and coagulation cascades, TNF signaling pathway are related with the onset of AS. Complement and coagulation cascades is related to immunoprotecive and regulatory function [51] and there are 36 formula targets engaged in it.

We also found that 13 drug targets were shared by more than 5 compounds. Among them, TNF, PTGS2, CYP1A1 (cytochrome P450 family 1 subfamily A member 1) seems to be hub targets because of the high frequency in herbs, showing a superposition effect. PPI network between drug targets and formula targets illustrated the underlying mechanism that why TCM can enhance the efficacy of western medicine. Furthermore, PPI networks between four datasets were constructed and hub proteins were selected in the network. The four data pieces formed a complex network, suggesting a close relationship between formula targets and AS. Results suggested that proteins such as JAK2, STAT3, HSP90AA1, TNF and PTEN are the key targets in PPI networks.

In all, the integrative investigation of the Y-Y-T targets in biological function helps us better understand how this formula treats the refractory disease AS. In this study, our systemic method is different from traditional single target research to reveal the therapeutic mechanism of Y-Y-T formula on AS treatment,

which provide an alternative way to investigate the TCM formulae.

## Materials and methods

### Clinical data collection

This retrospective study was carried out under the approval of Ethics committee of Zhejiang Chinese Medical University and the methods were carried out in accordance with the approved guidelines. Written informed consents were obtained from all patients before they participating in the study. We stuck to the basic principle of "freely given informed consent". Patients enrolled in the study range from 2015 to 2016 and fit with New York criteria revised in 1984. Patients' detailed information was collected either from outpatient medical records in The Second Affiliated Hospital of Zhejiang University of TCM or from the patients' follow-up visits. Those who had other diseases or suffering from liver and renal dysfunction were removed. Finally, 13 cases met the criteria. General information, serum markers (ESR and CRP) as well as indicators of disease progression were recorded. T-test was applied to analyze the otherness of indexes in different treating period.

### Collection of Y-Y-T data, AS disease proteins and AS drug targets

Ingredients of the 11 herbs in Y-Y-T and targets of those ingredients were gleaned from Traditional Chinese Medicine Integrated Database (TCMID; http://www.megabionet.org/tcmid/) [52]. Online Mendelian Inheritance in Man (OMIM; http://omim.org) [53] and The Genetic Association Database (GAD; http://geneticassociationdb.nih.gov) [54] were applied for the collection of AS disease proteins while AS drug targets were gathered from Drug Bank (http://www.drugbank.ca) [55].

### Microarray data processing of AS samples

72 expression matrix samples were download from GEO database. Among the 72 samples, 42 are from AS patient samples while 30 are from normal samples. Expression values were normalized by rma function in R. T-test with Benjamini-Hochberg correction for multiple comparisons were carried out to identify DEGs and adjust P-value less than 0.05 were defined as DEGs.

### GO enrichment analysis

GO and KEGG pathway enrichment analysis for Y-Y-T targets were undertaken through the online analytical tools DAVID Bioinformatics Resources 6.7 (http://david.abcc.ncifcrf.gov/) and topGO function in R.

**Target association network**

*Homo sapiens* protein-protein interaction data was extracted from InWeb_InBioMap, the most complete Human PPI database. Networks were visualized by Cytoscape 3.4.0.

**Acknowledgments**

This work was supported by National Basic Research Program of China (973 Program, No. 2014CB543001).

**References**

1 Duran, A. et al. Fecal calprotectin is associated with disease activity in patients with ankylosing spondylitis. Bosnian journal of basic medical sciences / Udruzenje basicnih mediciniskih znanosti = Association of Basic Medical Sciences 16, 71-74, doi:10.17305/bjbms.2016.752 (2016).

2 Osman, M. S. & Maksymowych, W. P. An update on the use of tumor necrosis factor alpha inhibitors in the treatment of ankylosing spondylitis. Expert review of clinical immunology, doi:10.1080/1744666X.2016.1218761 (2016).

3 Zhou, Y. Y., Lin, J. H., Huang, R. Y. & He, Y. T. Treatment of Ankylosing Spondylitis With a Bushen-Qiangdu-Zhilv Decoction: A Case Report With a 3-year Follow-up. Alternative therapies in health and medicine 22 Suppl 1, 36-40 (2016).

4 Machado, N. P. et al. Clinical characteristics and frequency of TLR4 polymorphisms in Brazilian patients with ankylosing spondylitis. Revista brasileira de reumatologia, doi:10.1016/j.rbr.2016.05.004 (2016).

5 Bae, J. M., Choo, J. Y., Kim, K. J. & Park, K. S. Association of inflammatory bowel disease with ankylosing spondylitis and rheumatoid arthritis: A nationwide population-based study. Modern rheumatology / the Japan Rheumatism Association, 1-6, doi:10.1080/14397595.2016.1211229 (2016).

6 Kim, M., Won, J. Y., Choi, S. Y., Ju, J. H. & Park, Y. H. Anti-TNFalpha treatment for HLA-B27 positive ankylosing spondylitis-related uveitis. American journal of ophthalmology, doi:10.1016/j.ajo.2016.07.016 (2016).

7 Mazouyes, A., Clay, M., Bernard, A. C., Gaudin, P. & Baillet, A. Efficacy of triple association methotrexate, sulfasalazine and hydroxychloroquine in early treatment of rheumatoid arthritis with insufficient response to methotrexate: Meta-analysis of randomized controlled trials. Joint, bone, spine : revue du rhumatisme, doi:10.1016/j.jbspin.2016.10.010 (2016).

8 Li, Q. et al. Kunxian capsules in the treatment of patients with ankylosing spondylitis: a randomized placebo-controlled clinical trial. Trials 17, 337, doi:10.1186/s13063-016-1438-6 (2016).

9 Liu, W., Zhang, D., Wu, Y. H. & Yang, H. J. [Efficacy Observation for Treating Ankylosing Spondylitis by Chinese Herbs and Recombinant Hu- man Tumor Necrosis Factor Receptor II-Antibody Fusion Protein]. Zhongguo Zhong xi yi jie he za zhi Zhongguo Zhongxiyi jiehe zazhi = Chinese journal of integrated traditional and Western medicine / Zhongguo Zhong xi yi jie he xue hui, Zhongguo Zhong yi yan jiu yuan zhu ban 36, 663-667 (2016).

10 Wang, C., Wang, G., Liu, H. & Hou, Y. L. Protective effect of bioactive compounds from Lonicera japonica Thunb. against H2O2-induced cytotoxicity using neonatal rat cardiomyocytes. Iranian journal of basic medical sciences 19, 97-105 (2016).

11 Wu, J. et al. Coptisine from Coptis chinensis inhibits production of inflammatory mediators in lipopolysaccharide-stimulated RAW 264.7 murine macrophage cells. European journal of pharmacology 780, 106-114, doi:10.1016/j.ejphar.2016.03.037 (2016).

12 Fatani, A. J. et al. Myrrh attenuates oxidative and inflammatory processes in acetic acid-induced ulcerative colitis. Experimental and therapeutic medicine 12, 730-738, doi:10.3892/etm.2016.3398 (2016).

13 Yang, J. et al. Anti-allodynic effect of intrathecal processed Aconitum jaluense is associated with the inhibition of microglial activation and P2X7 receptor expression in spinal cord. BMC complementary and alternative medicine 16, 214, doi:10.1186/s12906-016-1201-2 (2016).

14 Zhou, Q., Yu, D. H., Zhang, C., Liu, S. M. & Lu, F. Total saponins from Discorea nipponica ameliorate urate excretion in hyperuricemic mice. Planta medica 80, 1259-1268, doi:10.1055/s-0034-1383048 (2014).

15 Huang, L., Lv, Q., Liu, F., Shi, T. & Wen, C. A Systems Biology-Based Investigation into the Pharmacological Mechanisms of Sheng-ma-bie-jia-tang Acting on Systemic Lupus Erythematosus by Multi-Level Data Integration. Scientific reports 5, 16401, doi:10.1038/srep16401 (2015).

16 Huang, L., Lv, Q., Xie, D., Shi, T. & Wen, C. Deciphering the Potential Pharmaceutical Mechanism of Chinese Traditional Medicine (Gui-Zhi-Shao-Yao-Zhi-Mu) on Rheumatoid Arthritis. Scientific reports 6, 22602, doi:10.1038/srep22602 (2016).

17 Zhao, P. et al. Systems pharmacology-based approach for dissecting the active ingredients and potential targets of the Chinese herbal Bufei Jianpi formula for the treatment of COPD. International journal of chronic obstructive pulmonary disease 10, 2633-2656, doi:10.2147/COPD.S94043 (2015).

18 Zhang, Y. et al. A systems biology-based investigation into the pharmacological mechanisms of wu tou tang acting on rheumatoid arthritis by integrating network analysis. Evidence-based complementary and alternative medicine : eCAM 2013, 548498, doi:10.1155/2013/548498 (2013).

19 Guo, Y. et al. Therapeutic effect of dioscin on collagen-induced arthritis through reduction of Th1/Th2. International immunopharmacology 39, 79-83, doi:10.1016/j.intimp.2016.06.029 (2016).

20 Qi, M. et al. Dioscin alleviates lipopolysaccharide-induced inflammatory kidney injury via the microRNA let-7i/TLR4/MyD88 signaling pathway. Pharmacological research 111, 509-522, doi:10.1016/j.phrs.2016.07.016 (2016).

21 Meng, Q. F. et al. Astilbin ameliorates experimental autoimmune myasthenia gravis by decreased Th17 cytokines and up-regulated T regulatory cells. Journal of neuroimmunology 298, 138-145, doi:10.1016/j.jneuroim.2016.07.016 (2016).

22 Palocz, O., Paszti-Gere, E., Galfi, P. & Farkas, O. Chlorogenic Acid Combined with Lactobacillus plantarum 2142 Reduced LPS-Induced Intestinal Inflammation and Oxidative Stress in IPEC-J2 Cells. PloS one 11, e0166642, doi:10.1371/journal.pone.0166642 (2016).

23 Styrczewska, M. et al. Flax Fiber Hydrophobic Extract Inhibits Human Skin Cells Inflammation and Causes Remodeling of Extracellular Matrix and Wound Closure Activation. BioMed research international 2015, 862391, doi:10.1155/2015/862391 (2015).

24 Chenna Narendra, S., Chalise, J. P., Magnusson, M. & Uppugunduri, S. Local but Not Systemic Administration of Uridine Prevents Development of Antigen-Induced Arthritis. PloS one 10, e0141863, doi:10.1371/journal.pone.0141863 (2015).

25 Song, L. et al. Antibacterial activity of Pyrrosia petiolosa ethyl acetate extract against Staphylococcus aureus by decreasing hla and sea virulence genes. Natural product research, 1-4, doi:10.1080/14786419.2016.1244201 (2016).

26 Nalbant, S., Cagiltay, E., Sahan, B., Terekeci, H. M. & Oktenli, C. The vasoactive intestinal polypeptide (VIP) levels at the patients with ankylosing spondylitis and its association with inflammation markers. Rheumatology international 31, 1143-1146, doi:10.1007/s00296-010-1417-2 (2011).

27 Zhang, Y. et al. Implant-derived magnesium induces local neuronal production of CGRP to improve bone-fracture healing in rats. Nature medicine, doi:10.1038/nm.4162 (2016).

28 Yuan, T. L. et al. Serum Heme Oxygenase-1 and BMP-7 Are Potential Biomarkers for Bone Metabolism in Patients with Rheumatoid Arthritis and Ankylosing Spondylitis. BioMed research international 2016, 7870925, doi:10.1155/2016/7870925 (2016).

29 Li, Y. et al. Whole Genome Expression Profiling and Signal Pathway Screening of MSCs in Ankylosing Spondylitis. Stem cells international 2014, 913050, doi:10.1155/2014/913050 (2014).

30 Xu, L. et al. Changes in gene expression profiles of the hip joint ligament of patients with ankylosing spondylitis revealed by DNA chip. Clinical rheumatology 31, 1479-1491, doi:10.1007/s10067-012-2038-9 (2012).

31 Chen, C., Zhang, X. & Wang, Y. Analysis of JAK2 and STAT3 polymorphisms in patients with ankylosing spondylitis in Chinese Han population. Clinical immunology 136, 442-446, doi:10.1016/j.clim.2010.05.003 (2010).

32 Bonamichi-Santos, R. & Castells, M. Diagnoses and Management of Drug Hypersensitivity and Anaphylaxis in Cancer and Chronic Inflammatory Diseases: Reactions to Taxanes and Monoclonal Antibodies. Clinical reviews in allergy & immunology, doi:10.1007/s12016-016-8556-5 (2016).

33 Rice, J. W. et al. Small molecule inhibitors of Hsp90 potently affect inflammatory disease pathways and exhibit activity in models of rheumatoid arthritis. Arthritis and rheumatism 58, 3765-3775, doi:10.1002/art.24047 (2008).

34 Shalimar, A., Sharaf, I., Farah Wahida, I. & Ruszymah, B. H. Congenital insensitivity to pain with anhydrosis in a Malaysian family: a genetic analysis. Journal of orthopaedic surgery 15, 357-360 (2007).

35 Chen, Y. & Xia, R. G. Screening and functional microarray analysis of differentially expressed genes related to osteoporosis. Genetics and molecular research : GMR 13, 3228-3236, doi:10.4238/2014.April.25.8 (2014).

36 Zou, Y. C., Yang, X. W., Yuan, S. G., Zhang, P. & Li, Y. K. Celastrol inhibits prostaglandin E2-induced proliferation and osteogenic differentiation of fibroblasts isolated from ankylosing spondylitis hip tissues in vitro. Drug design, development and therapy 10, 933-948, doi:10.2147/DDDT.S97463 (2016).

37 Tragiannidis, A., Kyriakidis, I., Zundorf, I. & Groll, A. H. Invasive fungal infections in pediatric patients treated with tumor necrosis alpha (TNF-alpha) inhibitors. Mycoses, doi:10.1111/myc.12576 (2016).

38 Hung, Y. H. et al. IkBalpha promoter polymorphisms in patients with ankylosing spondylitis. Rheumatology international 30, 93-97, doi:10.1007/s00296-009-0923-6 (2009).

39 Zervou, M. I. et al. Lack of association of variants previously associated with anti-TNF medication response in rheumatoid arthritis patients: results from a homogeneous Greek population. PloS one 8, e74375, doi:10.1371/journal.pone.0074375 (2013).

40 Garrett, S. et al. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. The Journal of rheumatology 21, 2286-2291 (1994).

41 Chen, C. H. et al. The clinical usefulness of ESR, CRP, and disease duration in ankylosing spondylitis: the product of these acute-phase reactants and disease duration is associated with patient's poor physical mobility. Rheumatology international 35, 1263-1267, doi:10.1007/s00296-015-3214-4 (2015).

42 Rudwaleit, M. et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis and rheumatism 60, 717-727, doi:10.1002/art.24483 (2009).

43 Bredella, M. A., Steinbach, L. S., Morgan, S., Ward, M. & Davis, J. C. MRI of the sacroiliac joints in patients with moderate to severe ankylosing spondylitis. AJR. American journal of roentgenology 187, 1420-1426, doi:10.2214/AJR.05.1423 (2006).

44 Cock, I. E. & van Vuuren, S. F. The potential of selected South African plants with anti-Klebsiella activity for the treatment and prevention of ankylosing spondylitis. Inflammopharmacology 23, 21-35, doi:10.1007/s10787-014-0222-z (2015).

45 Carullo, G. et al. Quercetin and derivatives: useful tools in inflammation and pain management. Future medicinal chemistry 9, 79-93, doi:10.4155/fmc-2016-0186 (2017).

46 Reveille, J. D. Biomarkers for diagnosis, monitoring of progression, and treatment responses in ankylosing spondylitis and axial spondyloarthritis. Clinical rheumatology 34, 1009-1018, doi:10.1007/s10067-015-2949-3 (2015).

47 Briolay, A. et al. Autocrine stimulation of osteoblast activity by Wnt5a in response to TNF-alpha in human mesenchymal stem cells. Biochemical and biophysical research communications 430, 1072-1077, doi:10.1016/j.bbrc.2012.12.036 (2013).

48 Chidambara Murthy, K. N., Jayaprakasha, G. K. & Patil, B. S. The natural alkaloid berberine targets multiple pathways to induce cell death in cultured human colon cancer cells. European journal of pharmacology 688, 14-21, doi:10.1016/j.ejphar.2012.05.004 (2012).

49 Chen, F. L. et al. Berberine inhibits the expression of TNFalpha, MCP-1, and IL-6 in AcLDL-stimulated macrophages through PPARgamma pathway. Endocrine 33, 331-337, doi:10.1007/s12020-008-9089-3 (2008).

50 Kong, G. et al. Astilbin alleviates LPS-induced ARDS by suppressing MAPK signaling pathway and protecting pulmonary endothelial glycocalyx. International immunopharmacology 36, 51-58, doi:10.1016/j.intimp.2016.03.039 (2016).

51 Wiegner, R., Chakraborty, S. & Huber-Lang, M. Complement-coagulation crosstalk on cellular and artificial surfaces. Immunobiology 221, 1073-1079, doi:10.1016/j.imbio.2016.06.005 (2016).

52 Xue, R. C. et al. TCMID: traditional Chinese medicine integrative database for herb molecular mechanism analysis. Nucleic acids research 41, D1089-D1095, doi:10.1093/nar/gks1100 (2013).

53 Baxevanis, A. D. Searching Online Mendelian Inheritance in Man (OMIM) for information on genetic loci involved in human disease. Current protocols in human genetics / editorial board, Jonathan L. Haines ... [et al.] Chapter 9, Unit 9 13 11-10, doi:10.1002/0471142905.hg0913s73 (2012).

54 Becker, K. G., Barnes, K. C., Bright, T. J. & Wang, S. A. The genetic association database. Nature genetics 36, 431-432, doi:10.1038/ng0504-431 (2004).

55 Law, V. et al. DrugBank 4.0: shedding new light on drug metabolism. Nucleic acids research 42, D1091-1097, doi:10.1093/nar/gkt1068 (2014).

# 柠条锦鸡儿 *SHN* 基因的克隆及序列分析

红格日其其格，李高，王瑞刚*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail：pmb403@163.com

**摘要：**AP2/ERFs 是一个庞大的转录因子家族，对植物的生长发育过程中起着重要作用。根据已报道的蜡质合成相关的转录因子 SHN 基因序列设计引物，并利用 RACE 技术克隆获得柠条锦鸡儿的 AP2 转录因子编码基因 SHN 的全长 cDNA 序列。序列比对分析表明， SHN 序列长度为 999 bp，具有开放阅读框 651bp，编码 216 个氨基酸。预测该基因编码的蛋白分子量为 24.06KDa，等电点 5.62。氨基酸序列分析显示该序列具有 AP2 结构域，属于 AP2 超家族成员。系统进化分析表明该蛋白与大豆 SHN3 具有最高的同源性，将该基因命名为 CkSHN。

**关键词：**柠条锦鸡儿；SHN；RACE；序列分析

# 柠条锦鸡儿 *CkLEA6-1* 基因的克隆及功能分析

任方媛，杨杞，滑璐玢，王瑞刚，李国婧*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail：pmb403@163.com

**摘要：**在植物整个生命过程中，干旱、土壤盐碱化和冻害成为制约植物生长发育的重要逆境因子，造成农作物减产，导致生态环境日益恶化，因此植物在长期的进化过程中形成了不同的生理生化机制适应各种环境胁迫。胚胎发育晚期蛋白(Late embryogenesis abundant，LEA)便是植物在遭遇逆境时诱导合成的一系列功能蛋白，具有亲水性和热稳定性,有保护植物体、维持其生命代谢过程的功能。LEA 蛋白最早是在棉花发育后期的子叶中被发现的，随后，其他物种中的 LEA 蛋白也相继被报道。最初研究认为 LEA 蛋白是普遍存在于高等植物种子中的一类蛋白，当种子成熟脱水时可以保护植物的组织细胞免受伤害。后来经过大量研究表明，LEA 蛋白并不止在植物种子中大量表达，当植物受到干旱、低温、高盐以及植物激素脱落酸(ABA)等非生物胁迫时均能诱导 LEA 蛋白编码基因大量表达，因此大多 LEA 蛋白诱导表达无组织特异性。在细菌、真菌和一些无脊椎动物中也发现了 LEA 蛋白的存在，由此可见 LEA 蛋白并不是植物界所特有的。LEA 蛋白作为一类重要的抗逆蛋白越来越受关注，相关研究也日益增多。本研究从柠条锦鸡儿(*Caragana korshinskii* Kom.)中克隆得到一个 LEA 蛋白基因 *CkLEA6-1*。*CkLEA6-1* 基因的开放阅读框长为 255bp，编码 85 个氨基酸，且具有 LEA_6 保守结构域，因此命名该基因为 *CkLEA6-1*。实时荧光定量 PCR 检测发现 *CkLEA6-1* 的表达主要受冷胁迫诱导。根长实验发现在冷胁迫处理下 *CkLEA6-1* 基因过表达拟南芥幼苗的根长明显比野生型长，且幼苗鲜重略高于野生型。相对电导率实验发现在低温胁迫下，*CkLEA6-1* 基因过表达拟南芥质膜的相对电导率在常温、4℃时无明显差别，在-20℃时过表达株系较野生型略低，而且后续检测过表达株系的 MDA 含量后发现，4℃处理后过表达株系的 MDA 含量低于野生型。推测 *CkLEA*6-1 基因可能与柠条锦鸡儿响应逆境胁迫有关。

**关键词：**柠条锦鸡儿；*LEA* 基因；拟南芥；冷胁迫；植物抗逆性

# 中间锦鸡儿 *CiMYB60* 基因克隆及表达分析

刘建政，郝志霞，王瑞刚*，李国婧

内蒙古农业大学生命科学学院，呼和浩特 010018

*E-mail: pmb403@163.com

**摘要：**R2R3-MYB 转录因子在初生与次生代谢、生长发育及对生物与非生物胁迫的应答等各个方面具有重要作用。本研究以中间锦鸡儿为实验材料利用 PCR 技术分别以 cDNA 与 gDNA 为模板对 *CiMYB60* 基因进行了克隆。测序结果表明：*CiMYB60* 基因的开放阅读框为 1035 bp，编码 345 个氨基酸，其基因组 DNA 序列长度为 1485 bp，含有 2 个内含子和 3 个外显子。生物信息学分析显示：CiMYB60 所编码蛋白的 N 端包含 2 个 MYB 结构域，因此认为其属于典型的 R2R3-MYB 类蛋白。预测该蛋白分子量为 38.41 kDa，等电点为 5.95，蛋白整体上是亲水性的。利用染色体步移技术克隆到 *CiMYB60* 基因的 ATG 上游序列长度为 1848 bp，分析显示启动子序列中包含一些与光反应、组织特异性表达、激素和非生物胁迫相关的响应元件。利用实时荧光定量 PCR 技术对 *CiMYB60* 基因的表达进行分析，发现在脱水、NaCl 和 UV-B 处理下其表达量随胁迫处理时间的延长而降低。上述研究结果表明 *CiMYB60* 基因可能参与中间锦鸡儿对逆境胁迫的响应过程。

**关键词：**中间锦鸡儿；*CiMYB60*；表达分析；启动子克隆

# 中间锦鸡儿 *CiCHIL* 启动子的克隆及分析

杨天瑞，杨飞芸，韩晗，李国婧，王瑞刚*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail: pmb403@163.com

**摘要：**植物的次生代谢产物能够为植物的果实和花着色、有助于种子的形成、花粉的传播、适应外界逆境以及防御害虫的攻击。苯丙烷代谢途径是植物中一个重要的次生代谢途径，苯丙烷代谢通路上有几条主要的分枝，分枝下游产生了上千种化合物。查尔酮异构酶（CHI，chalcone isomerase)是苯丙烷通路上的一个关键酶，催化查尔酮到黄烷酮的反应，主要作用是帮助底物进行正确的分子内环化反应，生成的黄烷酮类化合物成为苯丙烷代谢途径下游产物的底物。从中间锦鸡儿干旱胁迫抑制性削减杂交文库中克隆得到一个 *CHI* 基因家族成员，利用巢式 PCR 反应对其进行基因上游启动子的克隆，利用启动子顺式元件分析网站分 *CiCHIL* 的启动子，预测启动子上游的顺式作用元件。

**关键词：**查尔酮异构酶；紫外顺式作用元件；中间锦鸡儿

# 中间锦鸡儿肉桂酰辅酶 A 还原酶基因的克隆和功能鉴定

刘坤，张子义，崔爽，韩晓东，李国婧，王瑞刚*

内蒙古农业大学生命科学学院，呼和浩特 010018

*E-mail: pmb403@163.com

**摘要：**肉桂酰辅酶 A 还原酶（cinnamoyl-CoA reductase, CCR）是催化木质素合成特异途径的第一个限速酶，对木质素的合成起重要作用。该研究从中间锦鸡儿中克隆得到两个 *CCR* 基因，分别命名为 *CiCCR2* 和 *CiCCR3*，其中 *CiCCR2* 基因开放阅读框长为 897bp，编码 299 个氨基酸，*CiCCR3* 基因开放阅读框为 966bp，编码 322 个氨基酸，并成功构建过表达载体 *p35S::CiCCR2* 和 *p35S::CiCCR3*，转化野生型拟南芥 Columbia-0，分别得到纯合体 8 个株系、9 个株系。检测过表达 *CiCCR2* 和 *CiCCR3* 基因的拟南芥，成熟时期和幼苗时期木质素含量均高于野生型拟南芥组织化学染色也验证了转基因株系木质素积累较野生型拟南芥多。并且 *CiCCR2* 和 *CiCCR3* 过表达植株的鲜重和干重比野生型增加。

**关键词：中间锦鸡儿；肉桂酰辅酶 A 还原酶；木质素；克隆**

# 中间锦鸡儿组织培养体系的建立

牛肖翠，潘雨，王瑞刚*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail: pmb403@163.com

**摘要：**中间锦鸡儿是我国西北部干旱荒漠地区生长的豆科锦鸡儿属灌木，具有极强的耐逆境能力，是一种重要的造林树种和饲用植物；还是研究植物抗逆的优良材料。为研究中间锦鸡儿与抗逆相关的基因及建立其遗传转化体系，建立中间锦鸡儿的无性繁殖体系十分必要。因而，本研究以中间锦鸡儿的茎段为外植体，愈伤组织诱导培养基为MS+$GA_3$ 0.15mg/L+6-BA 0.2mg/L+NAA 0.4mg/L+AC 0.1g/L+蔗糖25g/L+琼脂6g/L+肌醇0.1g/L，pH值5.8-6.0，愈伤诱导率为33.33%；丛生芽诱导培养基为MS+6-BA 1.0mg/L+NAA 0.2mg/L +蔗糖25g/L+琼脂6g/L+肌醇0.1g/L，pH值5.8-6.0，丛生芽诱导率为13.33%；生根培养基为MS+$GA_3$ 0.15mg/L+IAA 0.5mg/L+蔗糖20g/L+琼脂6g/L+肌醇0.1g/L，pH值5.8-6.0，生根率为60%；将经过生根培养后的小植株从培养瓶中取出，移栽成活率为100%。中间锦鸡儿无性繁殖体系初步建成。

**关键词：**中间锦鸡儿；组织培养；茎段；外植体

# 中间锦鸡儿 *CiNAC1* 基因促进拟南芥叶片的衰老

岳文冉，岳俊燕，于秀敏，李国婧*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail: pmb403@163.com

**摘要：**植物在受到干旱、高温、盐碱、冷、虫害等胁迫以后，细胞会受到不同程度的损伤。中间锦鸡儿具有耐寒、抗旱、耐盐碱和贫瘠等特点，具有良好的防风固沙和保持水土功能。NAC 转录因子家族是植物特有的、最大的转录因子家族之一，在植物生长发育过程中有重要的作用。本实验克隆得到中间锦鸡儿 *NAC1* 基因的全长 cDNA 序列，并转化野生型拟南芥得到转基因过表达纯合体株系，分析发现乙烯处理以后，*CiNAC1* 基因的过表达植株促进拟南芥叶片衰老，并且叶绿素降解相关基因与衰老相关基因在乙烯处理后表达量明显提高，转基因过表达纯合体株系的表达量高于野生型。因此，我们推断 *CiNAC1* 基因可能在乙烯诱导的叶片衰老过程发挥作用。

**关键词：**NAC 转录因子；*CiNAC1*；乙烯；叶片衰老

# 含聚异戊二烯植物的筛选及组分分析

王燕飞，曹美环，王瑞刚*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail: pmb403@163.com

**摘要：**天然橡胶（Natural rubber）是异戊二烯单元（$C_5H_8$）$_n$以 1,4-顺式构型连接在一起形成的生物高分子，因其有高弹性、防震、耐压等物理特性，被广泛的应用于医疗行业、工业、航天建造及生活中。本文通过裂解气相色谱法、红外光谱法以及质谱法对蒲公英和地锦草做了定性及组分分析研究。结果表明，红外光谱吸收峰都在聚异戊二烯特征吸收峰附近；用裂解气质联用法对其聚合物萃取，并对其聚合物做了定性及组分分析研究，都检测出其主体成分为聚异戊二烯。在对聚合物进行组分分析时发现两种植物的主成分大体相同，都含有植物醇及其异构体、环阿屯醇及烷基取代物和酯化物、一定量的饱和及不饱和的脂肪酸及酯化物、少量谷甾醇，而在蒲公英中含有香树素。

**关键词：**天然橡胶；定性；组分分析

# 从鹅绒藤中萃取聚异戊二烯及化学组分分析

于爽，曹美环，王瑞刚*

内蒙古农业大学生命科学学院，呼和浩特，010018

*E-mail: pmb403@163.com

**摘要：**天然橡胶（Natural Rubber，NR）是异戊二烯单元（$C_5H_8$）$_n$以 1,4-顺式构型连接在一起形成的生物高分子。本文利用裂解气质联用法对鹅绒藤进行聚合物萃取，并对其聚合物做了定性及组分分析研究，结果表明鹅绒藤中聚合物成分为植物醇及其异构体、环阿屯醇及烷基取代物和酯化物、一定量的饱和及不饱和的脂肪酸及酯化物、少量谷甾醇和羽扇豆醇，其聚合物主要成分为聚异戊二烯。本研究提出将鹅绒藤作为研究橡胶生物合成的模式植物。

**关键字：**天然橡胶；鹅绒藤；聚异戊二烯；模式植物

# CircPro：一个识别具有蛋白编码潜能性环状 RNA 的工具

孟宪文，陈铭*

浙江大学生命科学学院生物信息学系，杭州，3130058

*E-mail：mchen@zju.edu.cn

**摘要：**环状 RNA（circRNA）作为一类内源性 RNA，广泛存在于真核细胞中。它在不同生物学过程中不断涌现的作用表明，circRNA 是 RNA 世界重要的成员之一。大多数 circRNA 是通过 mRNA 前体反向剪切产生，形成了没有 5’帽子和 3’多聚腺苷酸尾巴的共价闭环的结构。此外，多数 circRNA 不与核糖体关联，因而过去 circRNA 被认为是非编码的。然而，最新的研究发现一些 circRNA 能够在体内产生蛋白质，而一发现极大地拓展了转录组和蛋白组的范围。为了深入研究 circRNA 翻译这一领域，我们开发了从高通量测序中识别具有蛋白编码潜能性 circRNA 的一个整合工具（CircPro，http://bis.zju.edu.cn/CircPro）。

**关键词：**环状 RNA; 核糖体；蛋白编码潜能性；高通量测序

# 通辽市及内蒙古民族大学场馆位置图



会务组联系电话：

【秘书组】陈宇杰：15149974350；陈琦：18868107759；沈小仙：13666606799

【统筹组】裴志利：13604753002；韩国志：13847573966；马玉露：15047527972

Novogene 诺禾致源


## 基因组学文章频发

诺禾致源成立六年来，专注于高通量测序技术在基因组测序领域的应用,并针对复杂基因组开发了全球领先的NOVOheter软件,已成功完成多个物种的全基因组从头组装工作,多篇基因组文章发表于Nature、Science等高水平期刊。

[1] Wang S, Zhang JB, Jiao WQ, et al. Scallop genome provides insights into evolution of bilaterian karyotype and development[J]. Nature Ecology & Evolution, 2017.
[2] Huang J, Zhang CM, Zhao X, . et. al. The Jujube Genome Provides Insights into Genome Evolution and the Domestication of Sweetness/Acidity Taste in Fruit Trees.2016.
[3] Li M, Chen L, Tian S, et al. comprehensive variation discovery and recovery of missing sequence in the pig genome using multiple de novo assemblies[J]. Genome Research, 2016.
[4] Zhang T, Hu Y, Jiang W, et al. Sequencing of allotetraploid cotton (Gossypium hirsutum L. acc. TM-1) provides a resource for fiber improvement. Nature Biotechnology, 2015, 33(5): 531-537.
[5] Zhou X, Wang B, Pan Q, et al. Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. Nature genetics, 2014, 46(12): 1303-1310.
[6] Li Y H, Zhou G, Ma J, et al. De novo assembly of soybean wild relatives for pan-genome analysis of diversity and agronomic traits[J]. Nature biotechnology, 2014, 32(10): 1045-1052.
[7] Li M, Tian S, Jin L, et al. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars[J]. Nature genetics, 2013, 45(12): 1431-1438.
[8] Qu Y, Zhao H, Han N, et al. Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau[J]. Nature communications, 2013, 4(4): 2071.

## 明星产品久经考验

**基因组组装** PacBio Sequel三代测序+新技术 10XGenomics / Bionano / Chicago / Hi-C+二代测序，不同测序策略组合，高指标完成基因组组装；

**组装注释类** 全面解析基因组遗传信息；

**组装变异检测** 全面挖掘各物种变异信息；

**功能基因挖掘** 挖掘与昆虫适应性及特殊性状相关基因；

**重测序产品** 遗传图谱构建、全基因组关联分析（GWAS）、群体进化分析等；

## 测序平台全球领先

诺禾致源在全球运行25台NovaSeq、30台HiSeq X、10台PacBio Sequel、10X Genomics等平台。

**北京诺禾致源科技股份有限公司**

电话：400-966-0802
邮箱：support@novogene.com
网址：www.novogene.com
地址：北京市海淀区金码大厦B座17/21层
天津市武清区/南京市江北新区

艮泰® GenTai 并行计算与模拟系统解决方案
Solutions for Parallel Computing and Simulation

*服务·方案·产品*

# 艮泰——您最忠诚的计算资源合作伙伴！

## 服务

- 模拟环境咨询、调研与测试
- 搭建、整合、调试和优化设施服务
- 大型计算项目调研与实施部署
- 并行虚拟文件系统
- 软/硬件系统售后服务
- 软件并行化与性能测试
- 计算资源租赁与托管
- 艮泰高性能计算基础架构服务
- Linux系统管理与MPI并行计算培训
- 软件开发项目合作

## 产品应用

- 科研计算模拟环境（艮泰第一、第二、第三系列产品）
- 艮泰定制化云计算方案（特色、平台）
- 艮泰独创硬双操作系统
- 集群计算模拟平台
- 全新四路静音图形站
- 一体机模拟平台（特色、设计与部署、优点、配置参数）
- 异构计算模拟平台
- 分层网络存储系统

## 科学计算模拟应用测试

- 流体力学--摩托车流体动力学模拟
- 第一性原理--sic本征值、态密度、波函数计算
- 地理信息系统--地图分析模拟
- 燃烧模拟--火灾仿真数值模拟
- 分子动力学--碳纳米管
- 计算生物信息--卫星烟草花叶病毒

## 产品配置

| | |
|---|---|
| 外观 | 机架式 / 塔式 |
| 处理器 | 可根据用户要求提供中端、高端服务器，可配置20、24、32、64、88及以上数量的物理计算核心 |
| 内存 | 借助高达2TB RECC DDR4内存，高效提升计算平台的运算速率 |
| 硬盘 | 可根据客户需求选择<100TB内置存储或>100TB的外置存储 |
| GPU | 支持Nvidia Q系列或AMD FireFro系列显卡以及Nvidia T系列或G系列GPU计算卡 |

## 艮泰优势

1. 并行环境的搭建及培训使用，节约时间和成本，解决更大规模的问题，迅捷便利，高效经济；
2. 根据客户的应用领域和具体研究方向，为客户提供相适应的软件，并且分析软件可行性和并行效果；
3. 根据客户的研究方向和规模，通过数据分析和理论指导，择优选取，增添部件，优化配置；
4. 提供计算资源租赁，无需硬件采购，无需自主搭建计算平台，可直接使用艮泰已搭建好的计算平台和软件，节约科研时间；针对客户时间，对客户进行简单操作培训，可快速上手进行操作，轻松完成科研计算。

## 加入我们

艮泰为您提供系统解决方案工程、系统集成工程师、软件售前工程师、售前测试工程师等职位，设立极具竞争力的薪酬福利体系，包括职责岗位津贴、团队领导岗位津贴、项目绩效提成和年终奖、期权奖励等，对于有特别贡献的员工将有机会进入公司管理层。工作地点：上海、西安、长沙、天津等地可选。

咨询热线 400-600-1251

网址：www.shgentai.com
邮箱：hpc@shgentai.com


# 内蒙古民族大学学校简介

内蒙古民族大学始建于 1958 年，是内蒙古自治区重点建设高校，内蒙古自治区人民政府和国家民族事务委员会共建高校，也是国家"十三五"中西部高校基础能力建设工程（二期）支持高校。

学校位于内蒙古自治区东部，坐落在被誉为"科尔沁草原明珠"的通辽市。现有教职工 1739 人，其中专任教师 1119 人，具有高级职称教师 570 人，具有博士学位教师 243 人，硕士学位教师 753 人。学校现有全日制在校学生 22591 人，其中普通本科学生 20819 人、普通专科学生 381 人、少数民族预科学生 119 人、硕士研究生 753 人、博士研究生 9 人、留学生 510 人。

学校设有 27 个教学单位、6 个教辅单位、1 个附属单位，开设本科专业 72 个，涵盖了经济学、法学、教育学、文学、历史学、理学、工学、农学、医学、管理学、艺术学等 11 个学科门类。有蒙药学博士学位点 1 个，硕士学位授权一级学科 11 个，涵盖了 73 个硕士学位授权二级学科，有专业硕士学位授权点 4 个。

近年来，学校获得多项荣誉。2009 年学校被评为全国“民族团结进步模范集体”，2012 年被评为“全区对外文化交流工作先进单位”和“自治区一级信誉等级涉外院校”，2014 年附属医院和蒙医药学院分别被评为“全国民族团结进步模范集体”、“全国教育系统先进集体”。

面向未来，学校秉承“博学明理、崇德至善”的校训，坚持“务实办学、创新办学、开放办学、特色办学”的理念，不断深化综合改革，促进科学发展、内涵发展和转型发展，努力把学校建设成为具有鲜明地方特点、民族特色和学科优势的高水平综合性民族大学。

## 招聘信息

内蒙古民族大学 2017 年度拟招聘工作人员 50 名。其中，博士研究生岗位 41 名，硕士研究生岗位 9 名。其中，定向招聘“蒙汉兼通”岗位 12 名。

内蒙古民族大学附属医院 2017 年度拟招聘工作人员 96 名，其中，博士研究生岗位 8 名，硕士研究生岗位 88 名 。其中，定向招聘“蒙汉兼通”岗位 27 名。

更多信息请参阅：
http://www.imun.edu.cn