

浙江省生物信息学学会

2017 年第二届“学会杯”生物信息学竞赛

浙江省生物信息学学会

Bioinformatics Society of Zhejiang Province

解码生命的奥秘

Bioinformatics: Decoding the mysteries of life

第二届“学会杯”生物信息学大赛



一、大赛背景

21 世纪是生物信息大数据的时代，生物信息学已是自然科学的核心领域之一。为了培养汇聚更多的生物信息学人才，激发师生和相关专业人员创新创业实战能力，加快生物信息学在科研及应用领域的普及。本学会联合浙江省各高校，科研单位及生物信息相关企业举办首届“学会杯”生物信息学大赛。

2016 年首届大赛吸引了全国 33 个单位的 27 支队伍参加，得到了 10 多个企业的热情赞助，初赛及决赛为来自不同地区和单位的生物信息学爱好者提供了相互学习切磋的绝佳机会，也为生物企业展示品牌、物色人才提供了一个平台。第二届比赛欢迎更多的队伍参加，让我们生物信息学人才一起学习共同进步，在 21 世纪这个生物的世纪里大展宏图！

二、大赛描述

大数据分析、精准医疗、胎儿健康早筛、突发传染病监测、超级细菌致病基因查找等诸多领域都离不开生物信息学的运用。此次大赛主要面向高校在校生（包括硕士生、博士生）和各界社会相关领域人才，旨在培养学生的数据分析、大数据动手能力，调动学生的学习兴趣；挖掘高校大数据人才，发现大数据优质项目，为企业输送数据分析与大数据人才，为学生创造就业机会，鼓励高校大数据创新创业。希望通过大赛培养发掘更多的生物

信息人才，让参赛队员接触科研和社会实践中的生物信息学问题，为科研和生物企业的发展培养提供更精锐的人才。

三、参赛规则

- (1) 组委会命题：各参赛队选择其中一题，通过优化的方案设计和程序运用解决问题。
- (2) 优秀自主项目提交、征集。
- (3) 参赛团队需由 1-5 人组成，设队长一名。不限制学科背景，不限地域。
- (4) 大赛评审委员会将根据参赛作品挑选 10 个队伍进入初赛，评定其中三个最强队进入决赛。
- (5) 参赛者提交的参赛作品必须是由参赛者独立完成的原创作品，不得抄袭，不得违反任何相关法律法规。
- (6) 大赛组委会所提供的数据集和平台仅限于本次大赛使用，不得用于其他任何目的，如若因此给数据提供方或平台提供方造成损失，由参赛选手负责。
- (7) 参赛作品的知识产权等相关权益归参赛队伍所有，但获奖作品赛后须在开源网站上公开发布，最终解释权归学会所有。

四、大赛题目

问题 1. 测序数据：<https://www.ncbi.nlm.nih.gov/sra/?term=SRP045129>

文章地址：<http://www.nature.com/nbt/journal/v33/n4/abs/nbt.3096.html>

利用公共数据库中的大豆重测序数据，使用全基因组关联分析（GWAS），找到部分人工驯化相关的区域，并进一步挖掘一些重要关联性状如株高、产油等。通过群体进化（WGS）研究参赛队可以描述实现方法，包括分析流程 pipeline 和脚本、使用的软件，以及预期结果

问题 2. 测序数据：<https://www.ncbi.nlm.nih.gov/sra/?term=SRP045767>

文章地址：<https://www.nature.com/ng/journal/v46/n11/full/ng.3117.html>

利用公共数据库中 360 份番茄重测序数据（包括 10 个野生型、333 个红果品种和 17 个商业化 F1 品种，以及果实大小的 BSA 测序 4 个样本包括 2 亲本和 2 子代混池）。使用全基因组关联分析（GWAS）、群体分层分析、基因组比对分析等方法，构建番茄变异组图谱，揭示野生小番茄向人工栽培大果番茄的驯化机制，找到控制大果实的驯化 QTL。参赛队可以描述实现方法，包括分析流程 pipeline、脚本、使用软件以及预期结果。

问题 3. 测序数据:

<http://trace.ddbj.nig.ac.jp/DRAsearch/submission?acc=DRA001730>

文章地址：<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0109999>

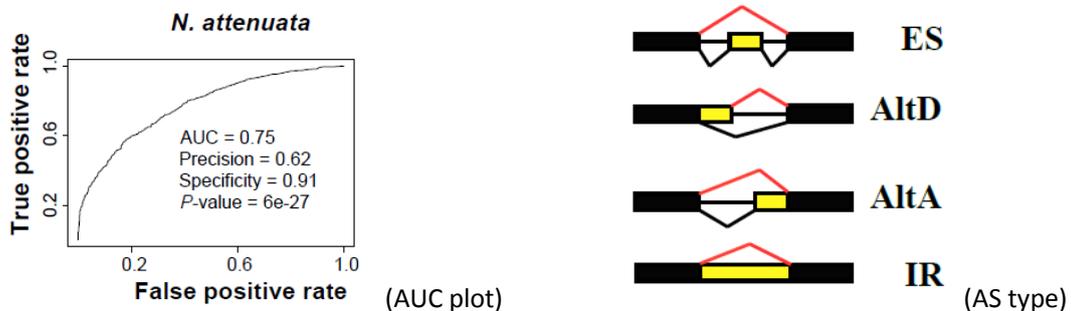
参考但不限于文章中的测序数据和思路，阐述细菌基因组完成图项目的建库、组装、基因组组分析、基因组注释以及比较基因组学方面的策略选择和理由。参赛队可描述实现方法，包括分析流程 pipeline、脚本、使用软件以及预期结果。

问题 4

可变剪切 (alternative splicing: AS) 是近年来比较热门的研究领域, 其将研究范畴从基因水平进一步扩展到了转录本水平, 而深度学习 (deep learning) 目前是研究模型, 统计规律里较热的技术。请分别查阅相关文献 (人类和拟南芥), 并从相应的 NCBI SRA 数据库里, 下载文章相关的 RNA-seq 数据, 从质控开始 (QC), 到最终检测出全基因组层面的所有 AS。之后查阅文献搜寻可能调控 AS 的相关因素 (e.g. [1][2])。利用深度学习的方法, 找出所导入因素 (在模型中每一个因素可视为一个参数) 对该类可变剪切的贡献度 (contribution), 判断所建模型的优劣 (AUC 图, 下图所示), 比较人类 (动物代表) 和拟南芥 (植物代表) 中的差异。只关注可变剪切的四大类 (下图所示)。

PS: deep learning 的模型构建可使用 R 的 H2O package

(<https://www.rdocumentation.org/packages/h2o/versions/3.10.4.6>), 也可使用其它软件, 任君挑选。



Reference:

1. Akerman M, Mandel-Gutfreund Y. Alternative splicing regulation at tandem 3' splice sites[J]. Nucleic Acids Research, 2006, 34(1):23.
2. Xia H, Bi J, Li Y. Identification of alternative 5'/3' splice sites based on the mechanism of splice site competition.[J]. Nucleic Acids Research, 2006, 34(21):6305-6313.

参赛选手可以选上面的任一题或者如下所列范围内的题目:

A、生物信息学**基础研究**项目, 如:

1. 基于 DaTo 生物信息学专业资源搜索系统(<http://bis.zju.edu.cn/DaTo2>)。
2. 生物数据一体化自动分析系统 (解决目前数据分析平台不同工具的整合、大数据上传和分析进度实时监测的问题)。
3. 一键出图插件或应用, 如自动生成染色体条带图、多层次调控网络可视化图等。
4. 基于机器学习, 通过足量的化验数据、医学图像影像等表型数据的训练, 实现智能辅助诊断, 开发相关应用软件。
5. 用生物信息学工具设计合成生物学模型或优化构建代谢途径。
6. 时序性周期性的特异表达基因的分群和聚类 (技术基础题)。
7. 基于 rule based 的 petri 网络通路设计 (备用)。

- B、**数据分析技巧**：基因组，转录组，非编码 RNA 拼接，注释分析。
C、**实用生物信息网站、软件、APP、科普游戏程序**。

五、大赛时间及报名

(1) 报名截止日期：2017 年 8 月 1 日

所有报名参赛队伍加入学会大赛 QQ 群下载报名表，需以 1-5 人为一小组填写报名表，统一发送至 ssping@zju.edu.cn 斯老师； huangfl@zju.edu.cn 黄老师，QQ: 7950200。组委会初步审核通过即报名成功。



学会杯大赛 QQ 群
219176940



微信公众号

(2) 方案提交截止日期：2017 年 9 月 10 日

参赛队伍需提交解决方案（软件源代码，PPT 讲解视频及其他支撑材料），由大赛组委会评审。

(3) 初赛评选日期：2017 年 9 月 30 日

筛选 10 支优秀参赛队伍，进入初赛答辩阶段，公平、公开、公正的公布初赛结果。
初赛地点：浙江大学紫金港校区生命科学学院生物信息中心实验室。

(4) 总决赛及颁奖仪式：2017 年 10 月 15 日（学会年会）

答辩现场根据参赛队的算法原理、历史成绩、评委打分，角逐出冠亚季军 3 只团队，复赛中余下的团队获得创新鼓励奖等企业奖项并颁发获奖证书。届时还将有学术界专家教授评委点评作品。

六、大赛奖项

- | | |
|--------|------|
| (1) 冠军 | 联川奖 |
| 季军 | 柏敖奖 |
| 最具潜力奖 | 厚泽奖 |
| 最具创意奖 | 遂真奖 |
| 最高效率奖 | 基因帮奖 |
| 最佳团队奖 | 高六博奖 |
| 最佳应用奖 | 可帮奖 |
| 特别奖 | 维迅奖 |

(2) 荣誉证书：大赛冠亚季军及其他奖项获得组委会颁发的荣誉证书

(3) 报名成功并提交解决方案者即可获得学会定制的精美纪念品一份。

七、大赛主办单位

浙江省生物信息学学会
浙江大学生命科学学院

八、大赛组委会联系方式：

学会官方网址：www.zjbioinformatics.org

斯老师 0571-88206134 转 8210 13805789825

黄老师 0571-88206593 13666652591（同微信号） QQ: 7950200

九、大赛组委会成员

主任 ： 陈铭

秘书处： 徐程、黄方亮、斯舒平、沈小仙

十、赞助单位



