



中国科学院北京基因组研究所（国家生物信息中心）
BEIJING INSTITUTE OF GENOMICS CHINESE ACADEMY OF SCIENCES / CHINA NATIONAL CENTER FOR BIOINFORMATION



国家基因组科学数据中心
National Genomics Data Center

科技项目数据汇交服务系统

Scientific Data Archive System, SDAS

国家生物信息中心
China National Center for Bioinformation

培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



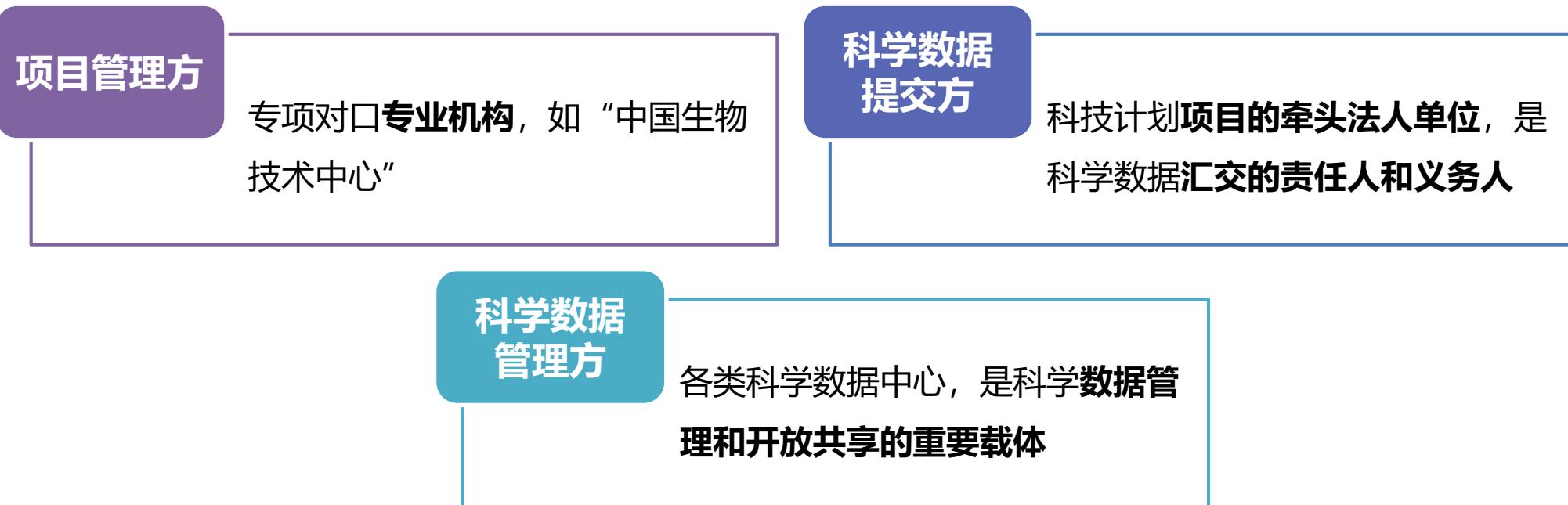
数据汇交政策依据

- 2018年3月17日，国务院办公厅印发《科学数据管理办法》（国办发〔2018〕17号），明确要求“政府预算资金资助的各级科技计划（专项、基金等）**项目所形成的科学数据**，应由**项目牵头单位汇交到相关科学数据中心**，各级计划（专项、基金等）项目管理部门应建立先汇交科学数据、再验收科技计划（专项、基金等）项目的机制。”
- 2018年12月14日，科技部办公厅印发《国家重点研发计划项目综合绩效评价工作规范（试行）》（国科办资〔2018〕107号），提出了由有关方面认定的**科学数据中心出具的汇交凭证**应作为项目综合绩效评价的必备材料。
- 2019年12月26日，科技部办公厅印发《科技计划项目科学数据汇交工作方案（试行）》（国科办基〔2019〕104号），进一步**明确了科技计划数据汇交的主体、职责、流程与方法**



数据汇交管理主体与职责

科学数据汇交由科技计划**项目管理方领导**，由**科学数据提交方**在**科学数据管理方**的支持下，协作完成。



数据汇交的内容

科学数据实体

- 项目形成的**原始数据**及基于原始数据或研究分析数据所形成的**完整数据库或数据文件**



- ✓ **自产多组学原始数据**
- ✓ **自产多组学的分析后数据**

科学数据描述信息

- 提供相应的**描述信息**，包括数据说明文档及过程资料等



- ✓ **多组学原始及分析后数据的项目、样本、实验、数据文件等描述元数据信息**

科学数据辅助工具

- 项目形成的用于科学数据处理、加工和分析的专门**辅助软件工具**等



- ✓ **辅助软件工具源代码等相关信息**
- ✓ **采购的工具软件不必汇交**



培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



系统介绍

科技项目数据汇交服务系统 (Scientific Data Archive System, SDAS) 作为国家基因组科学数据中心 (National Genomics Data Center, NGDC) 的汇交计划服务系统。

The screenshot shows the homepage of the SDAS system. At the top, there are navigation links for Databases, Tools, Standards, Publications, and About. Below the header, there's a main banner with icons representing DNA, an apple, a cell, a computer monitor, a test tube, and a petri dish. The banner text reads: "科学项目数据汇交服务系统 (Scientific Data Archive System, SDAS) 作为国家基因组科学数据中心 (National Genomics Data Center, NGDC) 的汇交计划服务系统。为生物领域国家重点专项提供汇交计划提交、实时数据归档情况查询、汇交证明出具等服务，帮助各级科技计划项目顺利验收。" Below the banner, there are five key statistics: 28 服务重点专项, 245 审核汇交计划, 101 出具汇交证明, 4522 归档数据集, and 2.63 总汇交数据量 (PB). At the bottom, there's a section titled "为什么选择我们?" (Why Choose Us?) with three items: "国际认可的数据平台" (International认可的数据平台), "人类遗传资源管理办公室指定备份平台" (Designated backup platform by the National Human Genetic Resources Management Office), and "多源异构数据兼容" (Compatibility with multi-source heterogeneous data).

系统定位:

- ✓ 服务项目层

主要功能:

- ✓ 汇交计划提交
- ✓ 汇交进度实时查询
- ✓ 汇交证明出具



系统介绍

科技项目数据汇交服务系统 (Scientific Data Archive System, SDAS) 作为国家基因组科学数据中心 (National Genomics Data Center, NGDC) 的汇交计划服务系统。

The screenshot shows the homepage of the SDAS system. At the top, there are navigation links for Databases, Tools, Standards, Publications, and About. Below the header is a banner for the National Genomics Data Center (NGDC). The main content area features a hexagonal grid icon representing biological data types like DNA, RNA, and proteins. Below this are five key statistics: 28 service focus projects, 245 reviewed submission plans, 101 issued data exchange certificates, 4522 archived datasets, and a total data volume of 2.63 PB. A section titled "为什么选择我们?" (Why Choose Us?) highlights three features: international recognition, designated backup for human genetic resources, and multi-source heterogeneous data compatibility.

服务重点专项	审核汇交计划	出具汇交证明	归档数据集	总汇交数据量 (PB)
28	245	101	4522	2.63

为什么选择我们?

- 国际认可的数据平台
- 人类遗传资源管理办公室指定备份平台
- 多源异构数据兼容

我们的优势

- ✓ 数据库国际认可度高
- ✓ 人遗资源数据指定备份平台
- ✓ 兼容多源异构数据



系统使用流程



系统使用流程

注册并提交汇交计划

1. 项目负责或联系人注册
2. 填写项目基本信息
3. 组织制定并提交计划

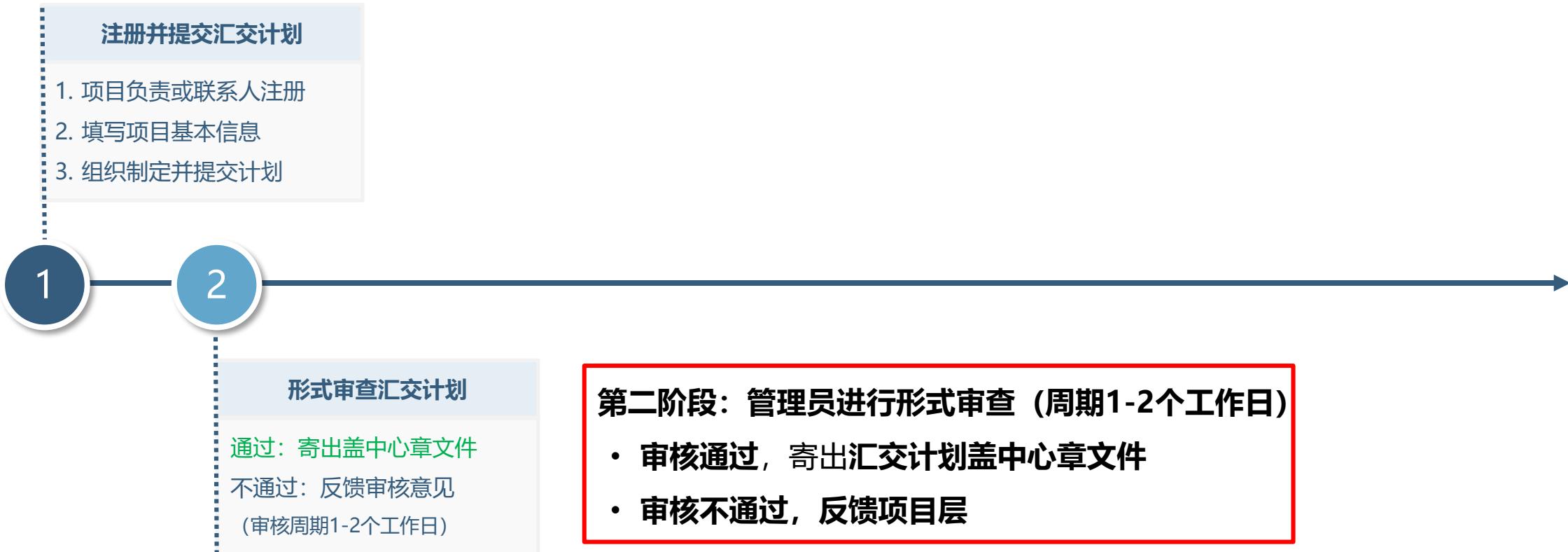
1

第一阶段：

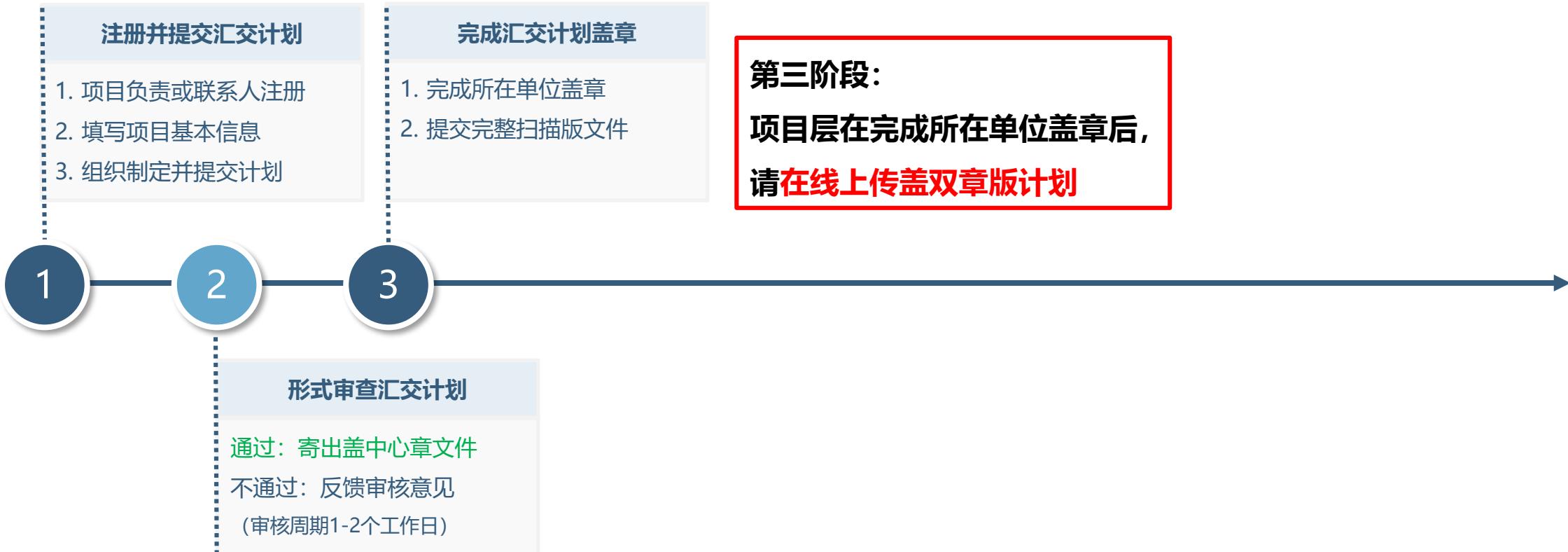
1. 使用项目负责人或联系人账号注册
2. 组织课题层制定并最终形成一份项目汇交计划，并在线提交



系统使用流程



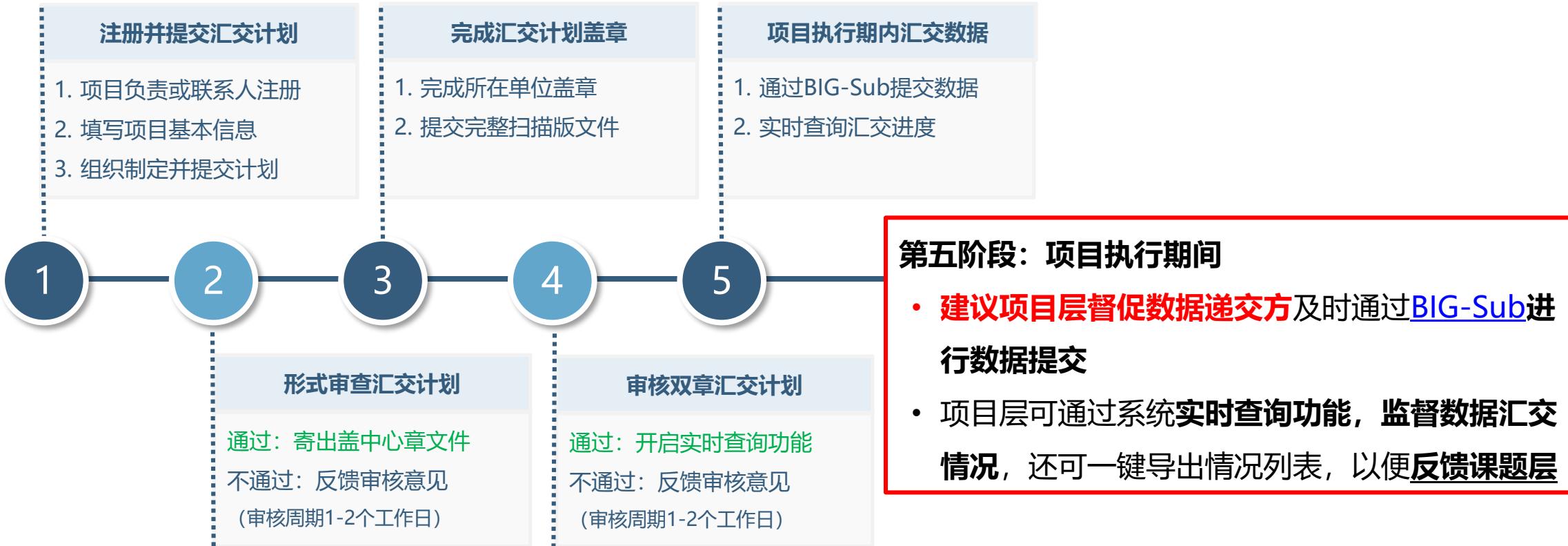
系统使用流程



系统使用流程



系统使用流程



系统使用流程



系统使用流程



培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



系统账号注册

- 建议项目层使用项目负责人/联系人邮箱通过中心单点登录系统 (BIG SSO) 完成账号注册。具体流程如下：
 - 申请注册账号：**点击SDAS系统主页菜单栏中的“注册”：

The screenshot shows the homepage of the SDAS system. At the top, there is a navigation bar with links for '主页' (Home), '提交' (Submit), '帮助' (Help), '注册' (Register) which is highlighted with a red box, and '登录' (Login). Below the navigation bar, there is a banner with the text '点击“注册”进行用户注册' (Click 'Register' to perform user registration). A callout box contains the text '填写信息提交后，一定记得完成激活！！！' (After filling in the information and submitting, remember to complete the activation!!!). Another callout box contains the text '说明文档： [SDAS系统使用指南](#)' (Documentation: [SDAS system usage guide](#)). At the bottom left, there are two buttons: '汇交计划提交入口' (Submission entry for collection plans) and '科学数据提交入口' (Submission entry for scientific data). On the left side, there is a brief introduction to the SDAS system: '科学项目数据汇交服务系统' (Science project data submission service system) and '科学项目数据汇交服务系统 (SDAS) 作为国家基因组科学数据中心 (National Genomics Data Center, NGDC) 的汇交计划服务系统。为生物领域国家重点专项项目提供汇交计划提交、实时数据归档情况查询、汇交证明出具等服务，帮助各级科技计划项目顺利验收。' (The SDAS system is a submission plan service system for the National Genomics Data Center (NGDC). It provides services such as submission plan submission, real-time data archiving status query, and submission certificate issuance for key national projects in the biological field, helping various science and technology plan projects to pass inspection smoothly.).



系统账号注册-常见问题

- 忘记激活账号 或 忘记登陆密码：打开登录页面或使用此链接 <https://ngdc.cncb.ac.cn/sso/login>



制定&修改汇交计划

- 项目层可通过“[汇交计划提交入口](#)”进入提交系统，点击“新建提交”开始提交

国家基因组科学数据中心
National Genomics Data Center

科学项目数据汇交服务系统

科学项目数据汇交服务系统 (Scientific Data Archive System, SDAS) 作为国家基因组科学数据中心 (National Genomics Data Center, NGDC) 的汇交计划服务系统。为生物领域国家重点专项项目提供汇交计划提交、实时数据归档情况查询、汇交证明出具等服务，帮助各级科技计划项目顺利验收。

汇交计划提交入口 科学数据提交入口

科技项目数据汇交服务系统

主页 / 个人中心

汇交计划制定
-- 点击【新建提交】开始在线递交汇交计划，具体流程详见【项目汇交计划制定】。提交后请耐心等待审核，审核反馈周期1-2个工作日。

在线查询&出具汇交证明
-- 计划审核通过并完成盖章版汇交计划上传后，即可开启数据归档情况在线查询和汇交证明申请出具功能，具体流程详见【项目汇交证明出具】。

新建提交 点击“新建提交”开始提交项目汇交计划



制定&修改汇交计划

- 第一步 填写项目相关信息：主要用于收集项目负责人、项目联系人、项目基本信息等内容

新建提交

1 基本信息 > 2 汇交计划制定 > 3 汇交协议签署 > 4 信息预览 > 5 提交完成

项目负责人信息
项目负责人姓名（中文）将出现在后续汇交证明中，请务必保证填写内容真实无误

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 联系电话（座机/手机）
座机格式：区号+号码，如010-84097340

项目联系人信息
项目联系人负责项目提交中的沟通工作，请确保联系电话畅通

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 手机号码
科技部信息管理要求，请务必提供

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中，请务必保证填写内容真实无误
鼠标移至问号 可查看表中各部分内容在项目书中的标注位置

* 项目资助号 资助号与申请号必须填写
 为正式资助号，填写项目申请编号

* 项目名称 项目申报指南方向
非必填内容

* 项目类别 * 专业机构 * 专项名称

* 起始时间 * 结束时间

* 项目承担单位全称（中文） * 项目承担单位机构性质

课题承担单位名称（中文）

保存并进入下一步

项目负责人信息
项目负责人姓名（中文）将出现在后续汇交证明中，请务必保证填写内容真实无误

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 联系电话（座机/手机）
座机格式：区号+号码，如010-84097340

项目联系人信息
项目联系人负责项目提交中的沟通工作，请确保联系电话畅通

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 手机号码
科技部信息管理要求，请务必提供



制定&修改汇交计划

- 第一步 填写项目相关信息：主要用于收集项目负责人、项目联系人、项目基本信息等内容

1 基本信息 > 2 汇交计划制定 > 3 汇交协议签署 > 4 信息预览 > 5 提交完成

新建提交

项目负责人信息
项目负责人姓名(中文)将出现在后续汇交证明中,请务必保证填写内容真实无误

* 姓名(中文) * 职称 * 职务
如没有可填写“无”

* 邮箱 * 联系电话(座机/手机)
座机格式: 区号+号码, 如010-84097340

项目联系人信息
项目联系人负责项目提交中的沟通工作,请确保联系电话畅通

* 姓名(中文) * 职称 * 职务
如没有可填写“无”

* 邮箱 * 手机号码
科技部信息管理要求,请务必提供

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中,请务必保证填写内容真实无误
鼠标移至“问号”可查看表中各项内容在项目书中的标注位置

* 项目资助号
资助号与申请号必须择一填写
 暂无正式资助号, 填写项目申请编号

* 项目名称
项目申报指南方向
非必填内容

* 项目类别
专业机构
专项名称

* 起始时间
结束时间

* 项目承担单位全称(中文)
项目承担单位机构性质

课题承担单位名称(中文)

课题承担单位名称(中文)

保存并进入下一步

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中,请务必保证填写内容真实无误
鼠标移至“问号”可查看表中各项内容在项目书中的标注位置

项目资助号
资助号与申请号必须择一填写
 暂无正式资助号, 填写项目申请编号

项目名称
项目申报指南方向
非必填内容

项目类别
专业机构
专项名称

起始时间
结束时间

项目承担单位全称(中文)
项目承担单位机构性质

课题承担单位名称(中文)



制定&修改汇交计划

- 第一步 填写项目相关信息：主要用于收集项目负责人、项目联系人、项目基本信息等内容

1 基本信息 > 2 汇交计划制定 > 3 汇交协议签署 > 4 信息预览 > 5 提交完成

新建提交

项目负责人信息
项目负责人姓名（中文）将出现在后续汇交证明中，请务必保证填写内容真实无误

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 联系电话（座机/手机）
座机格式：区号+号码，如010-84097340

项目联系人信息
项目联系人负责项目提交中的沟通工作，请确保联系电话畅通

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 手机号码
科技部信息管理要求，请务必提供

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中，请务必保证填写内容真实无误
鼠标移动至“问号”可查看表中各项内容在项目书中的标注位置

* 项目资助号
资助号与申请号必须择一填写
 暂无正式资助号，填写项目申请编号

* 项目名称
项目申报指南方向
暂无正式资助号，填写项目申请编号

* 项目类别
 暂无正式资助号，填写项目申请编号

* 起始时间
结束时间

* 项目承担单位全称（中文）
课题承担单位名称（中文）

* 项目承担单位机构性质

保存并进入下一步

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中，请务必保证填写内容真实无误
鼠标移动至“问号”可查看表中各项内容在项目书中的标注位置

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中，请务必保证填写内容真实无误
鼠标移动至“问号”可查看表中各项内容在项目书中的标注位置

* 项目资助号
资助号与申请号必须择一填写
 暂无正式资助号，填写项目申请编号

* 项目名称
项目申报指南方向
暂无正式资助号，填写项目申请编号

* 项目类别
 暂无正式资助号，填写项目申请编号

* 起始时间
结束时间

* 项目承担单位全称（中文）
课题承担单位名称（中文）

* 项目承担单位机构性质



制定&修改汇交计划

- 第一步 填写项目相关信息：主要用于收集项目负责人、项目联系人、项目基本信息等内容

新建提交

1 基本信息 > 2 汇交计划制定 > 3 汇交协议签署 > 4 信息预览 > 5 提交完成

项目负责人信息
项目负责人姓名（中文）将出现在后续汇交证明中，请务必保证填写内容真实无误

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 联系电话（座机/手机）
座机格式：区号+号码，如010-84097340

项目联系人信息
项目联系人负责项目提交中的沟通工作，请确保联系电话畅通

* 姓名（中文） * 职称 * 职务
如没有可填写“无”

* 邮箱 * 手机号码
科技部信息管理要求，请务必提供

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中，请务必保证填写内容真实无误
鼠标移动至“问号”可查看表中各项内容在项目书中的标注位置

* 项目资助号
资助号与申请号必须择一填写
 暂无正式资助号

* 项目名称
项目申报指南方向
非必填内容

* 项目类别

* 起始时间
结束时间

* 项目承担单位全称（中文）
课题承担单位名称（中文）

保存并进入下一步

项目基本信息
项目资助号和项目名称将出现在后续汇交证明中，请务必保证填写内容真实无误
鼠标移动至“问号”可查看表中各项内容在项目书中的标注位置

项目编号：2017YFC120**00 密 级：公开

国家重点研发计划
项目任务书

项目名称：国家生物信息平台支撑技术
所属专项：生物安全关键技术研发
指南方向：国家生物信息平台支撑技术
推荐单位：工业和信息化部
专业机构：中国生物技术发展中心

项目牵头承担单位：***** (公章)
项目负责人：***
执行期限：2017年11月至2020年11月



制定&修改汇交计划

- 第二步 组织课题层共同制定计划：请项目层组织课题层制定并最终形成一份项目汇交计划。更多帮助，如填写示例，详见《国家基因组科学数据中心数据汇交指南》

SDAS000322

✓ 接收的数据类别

- 白质组学质谱及其分析后数据, 包括基因组、蛋白组、代谢组 (包括脂质组)
- 生物医学试剂盒或商业数据
- 自主研发的软件工具或数据库系统
- 其它的结果性、有共享价值的科学数据

✗ 不接收的数据类别

- 整合的公共数据
- 商用软件工具
- 实验的原始“一手”实验记录, 数据库说明文档、软件著作权、专利及文章等材料

帮助

常见问题 (FAQ)

汇交数据类型

* 请选择本项目提交自产数据类型 (可多选)

+ 组学原始测序下机数据 + 全基因组数据 + 基因型变异数据 + 基因表达分析数据 + 芯片数据 + 代谢组数据
+ 蛋白高通量测序数据 + 蛋白结构数据 + 细胞共培养数据 + 软件工具 + 数据库系统 + 其他数据

请至少选择一种数据类型

汇交计划文件

请选择Word版汇交计划，并单击上传按钮，上传已选择的文件

汇交计划 (word版)

附件材料 (非必填)

您可以上传项目任务书作为数据汇交计划的辅助材料 (非必须提供)
请单击此处选择文件后，点击右侧按钮上传

保存 下一步 返回上一步



制定&修改汇交计划

- 第二步 组织课题层共同制定计划：请项目层组织课题层制定并最终形成一份项目汇交计划。更多帮助，如填写示例，详见[《国家基因组科学数据中心数据汇交指南》](#)

The screenshot shows the 'Huijiao Plan Preparation' step (Step 2) of the SDAS system. The top navigation bar shows steps 1 through 5: 基本信息 (Basic Information), 汇jiao计划制定 (Huijiao Plan Preparation), 汇jiao协议签署 (Protocol Signing), 信息预览 (Preview), and 提交完成 (Submission Completed). Step 2 is highlighted in red.

The main content area displays the 'SDAS000322:' identifier. Below it, there are two sections: '接收的数据类别' (Accepted Data Categories) and '不接收的数据类别' (Non-Accepted Data Categories).

接收的数据类别 (Accepted Data Categories):

- ✓ 接收的数据类别
 - 自产多组学原始及其分析后数据，包括基因组、蛋白组、代谢组（包括脂质组）等
 - 脑科学领域核磁共振成像数据、电生理等数据
 - 自主研发的软件工具或数据库系统
 - 其它的结果性、有共享价值的科学数据

不接收的数据类别 (Non-Accepted Data Categories):

- ✗ 不接收的数据类别
 - 整合的公共数据
 - 商用软件工具
 - 实验的原始“一手”实验记录，数据库说明文档、软件著作权、专利及文章等材料

Below these sections, there are help links: 帮助 (Help) and 常见问题 (FAQ). At the bottom, there is a section for 'Huijiao Data Types' with checkboxes for '组学原始测序下机数据' (Genomic raw sequencing data), '全基因组组装数据' (Whole genome assembly data), '基因组变异数据' (Genome variation data), '基因组二级分析数据' (Genome secondary analysis data), and '芯片数据' (Chip data).

A large red box highlights the 'Accepted Data Categories' and 'Non-Accepted Data Categories' sections, emphasizing the differences between them.



制定&修改汇交计划

- 第二步 组织课题层共同制定计划：请项目层组织课题层制定并最终形成一份项目汇交计划。更多帮助，如填写示例，详见[《国家基因组科学数据中心数据汇交指南》](#)

汇交数据类型

* 请选出本项目提交自产数据类型（可多选）

+ 组学原始测序下机数据 + 全基因组组装数据 + 基因组变异数据 + 基因组二级分析数据 + 芯片数据 + 代谢组数据
+ 蛋白高通量测序数据 + 蛋白结构数据 + 核磁共振成像数据 + 软件工具 + 数据库系统 + 其他数据

请至少选择一种数据类型

汇交计划文件

请下载科学数据汇交计划提交模板文件《汇交计划通用模板》
更多帮助，如填写示例，请查看《国家基因组科学数据中心数

* 汇交计划（word版）

请选择word版汇交计划，并单击“上传”按钮，上传已选择的

附件材料（非必填）

您可以上传项目任务书作为数据汇交计划的辅助材料（非必须提供
请单击此处选择文件后，点击右侧按钮上传

汇交数据类型

* 请选出本项目提交自产数据类型（可多选）

✓ 组学原始测序下机数据 + 全基因组组装数据 + 基因组变异数据 + 基因组二级分析数据 + 芯片数据 + 代谢组数据
+ 蛋白高通量测序数据 + 蛋白结构数据 + 核磁共振成像数据 + 软件工具 + 数据库系统 + 其他数据

**用于辅助审核人员了解您
项目涉及的数据类别**



制定&修改汇交计划

- **第二步 组织课题层共同制定计划：**请项目层组织课题层制定并最终形成一份项目汇交计划。更多帮助，如填写示例，详见[《国家基因组科学数据中心数据汇交指南》](#)

w 汇交计划文件
请下载科学数据汇交计划提交模板文件《汇交计划通用模板》，完成填写并检查无误后上传（仅接收word版）
更多帮助，如填写示例，请查看《国家基因组科学数据中心数据汇交指南》

* 汇交计划 (word版)

上传

● 请选择word版汇交计划，并单击“上传”按钮，上传已选择的文件

w 附件材料 (非必填)
您可以上传项目任务书作为数据汇交计划的辅助材料（非必须）
请单击此处选择文件后，点击右侧按钮上传

保存 **下一步** **返回上一步**

w 汇交计划文件
请下载科学数据汇交计划提交模板文件《汇交计划通用模板》，完成填写并检查无误后上传（仅接收word版）
更多帮助，如填写示例，请查看《国家基因组科学数据中心数据汇交指南》

* 汇交计划 (word版)

上传

● 请选择word版汇交计划，并单击“上传”按钮，上传已选择的文件



制定&修改汇交计划

- **第三步 在线签订汇交协议：**请项目层务必仔细阅读汇交协议，同意并遵守所有条款。此外，请正确填写汇交计划盖章文件的邮寄信息。

The screenshot shows the 'NGDC Data Submission Agreement' step (Step 3) of a submission plan process. The top navigation bar shows five steps: 1. 基本信息, 2. 汇交计划制定, 3. 汇交协议签署 (highlighted in red), 4. 信息预览, and 5. 提交完成.

The main content area displays the 'NGDC Data Submission Agreement' document, which includes the following text:

SDAS000322:

NGDC Data Submission Agreement

1.承诺本项目所有数据自愿纳入国家基因组科学数据中心（下文简称“中心”）平台管理，遵循中心对汇交的科学数据进行分类、分级存储和管理的各项规范，不会擅自修改和删除汇交的科学数据；
2.承诺本项目遵守中心相关技术标准规范，保障汇交的数据真实、准确、完整、有效，无学术不端和学术失范行为，并承担因数据不真实等问题带来的一切后果和法律责任；
3.承诺本项目汇交数据遵循国际知识产权法（著作权法）和中国著作权法的有关规定，不存在对汇交的科学数据不具自主知识产权或有剽窃行为；
4.承诺组织项目课题承担单位完成数据汇交，确保后续汇交工作的顺利展开与完成。

我已经阅读了上述汇交计划生成免责声明，同意并遵守以上所有条款

邮寄地址
盖章版汇交计划将会按此地址邮寄，请确保真实有效

* 收件人姓名 * 手机号码

请填写收件人姓名 请填写收件人手机号码

* 收件地址
例：北京市朝阳区北辰西路1号院104号楼
请填写收件地址

备注
如希望自取，请注明

A large red box highlights the 'NGDC Data Submission Agreement' text and the checkbox below it. Another large red box highlights the entire 'NGDC Data Submission Agreement' section, including the text and the checkbox.



制定&修改汇交计划

- 第三步 在线签订汇交协议：请项目层务必仔细阅读汇交协议，同意并遵守所有条款。此外，请正确填写汇交计划盖章文件的邮寄信息。

The screenshot shows the third step of the submission plan process: 'Holding Submission Agreement'. The top navigation bar includes steps 1 (Basic Information), 2 (Submission Plan Preparation), 3 (Holding Submission Agreement), 4 (Preview Information), and 5 (Submission Completed). The main content area displays the title 'SDAS000322: 区域特异性神经干细胞的获取以及功能特性和应用的研究' and the 'NGDC Data Submission Agreement' section. This section contains a list of 4 promises and a checkbox for accepting the agreement. Below this is a 'Shipping Address' section with fields for收件人姓名 (Recipient Name) and 手机号码 (Mobile Number), both marked with red asterisks indicating required fields. There are also fields for 'Address Example' and 'Recipient Address'. At the bottom are 'Save and Submit' and 'Back' buttons.

如果希望自取，需在备注中写明，
否则默认快递到付寄出。



制定&修改汇交计划

- **第四步 信息预览，完成提交：**请项目层务必认真核对后，再点击提交。管理员反馈审核意见前，您可随时修改。

SDAS000270: 新冠病毒基因组序列的演化分析

项目基本信息

项目名称: 项目类别: 项目类型

项目负责人姓名: 职称/职务: 联系电话:

项目联系人姓名: 职称/职务: 联系电话:

项目汇交计划

提交时间: 2022-12-02 08:50:30

提交人: 手机号码:

提交 **返回上一步**

1. 点击“提交”，完成填写，跳回列表界面
2. 点击“返回上一步”，进行修改



制定&修改汇交计划

- **第四步 信息预览，完成提交：**请项目层务必认真核对后，再点击提交。管理员反馈审核意见前，您可随时修改。

该截图展示了系统提交界面的组成部分：

- 顶部导航：**显示“主页 / 个人中心”。
- 功能模块：**包含“汇交计划制定”、“在线查询&出具汇交证明”以及“新建提交”按钮。
- 状态提示：**显示“个人项目列表 (您有1信息未读)”，并包含红色框标注的“收到系统反馈信息，请点击“信息反馈”进行查看”。
- 详细信息区：**显示“系统编号: SDAS000322”、“项目编号: [REDACTED]”、“项目名称: [REDACTED]”和“项目负责人: [REDACTED]”。
- 流程说明：**显示“① 制定&修改汇交计划”、“② 上传盖双章版汇交计划”和“③ 在线查询&出具证明”。
- 状态说明：**显示“【状态说明】汇交计划待审核：请等待系统审核（审核周期1-2工作日），管理员反馈审核意见前，可随时修改”。
- 操作按钮：**包含一个带有“信息反馈”按钮的红色框。

**完成提交后，请等待管理员进行审核
(审核周期1-2工作日)**

注意：管理员反馈审核意见前，您可随时修改。请尽量避免多次修改

显示状态和相关信息提示

查看审核意见和历史记录信息



SDAS000183: 项目1

[联系管理员](#)

联系我们

信息反馈页面

项目动态

● 管理员 2022-05-19 11:17:35	【审核完毕】审核通过 反馈: 1) 审核通过: 您上传的文件没有问题。后续您们可通过中心统一提交入口 (BIG-Sub: https://ngdc.cncb.ac.cn/gsub/) 完成各类型数据提交。归档数据实时查询与汇交证明出具流程详见视频 "XXXXXX" 。
● 用户 Zhang Sisi 2022-05-19 11:17:26	【SDAS000183系统提交】 null: 项目1 上传盖双章版汇交计划
● 管理员 2022-05-19 11:16:54	【SDAS000183上传汇交计划】 null: 项目1 盖单章版计划已寄出 备注: 汇交计划盖章页已使用顺丰寄出, 请您注意查收。
● 管理员 2022-05-19 11:16:35	【审核完毕】审核通过 反馈: 汇交计划审核通过, 我们将在1-2个工作日内为您寄出一式四份的纸质版盖章页; 项目承担单位需寄送两份给所对应专业机构的专项办公室, 以便专业机构审核归档, 具体寄送方式及细节请咨询对应专项; 当您们完成项目负责人和项目承担单位盖章后, 请点击 "上传盖双章版汇交计划" 完成盖双章扫描版汇交计划 (PDF) 文件。 待盖双章版计划审核通过后, 即可开启数据汇交情况实时查询和在线申请出具汇交证明。
● 用户 Zhang Sisi 2022-05-19 11:16:17	【SDAS000183系统提交】 null: 项目1 修改原因: null 修改汇交计划后重新提交
● 管理员 2022-05-11 10:11:01	【审核完毕】审核不通过 反馈: 附件: SDAS系统使用指南.pdf

下载审核反馈附件

- 查看所有重要节点反馈信息和材料
- 通过邮件联系管理员



汇交计划审核情况反馈

- 如果审核未通过：请根据**[信息反馈]**中的建议进行修改
 - 点击**[制定&修改汇交计划]**进入浏览页面



汇交计划审核情况反馈

- 点击[修改]进入汇交计划修改页面。此外，您可通过[更新项目信息]完成相关信息更新，包括更换联系人；

The screenshot shows a project submission page with the following details:

项目基本信息	
项目资助号	SDAS000322
项目申报指南方向	中国生物技术发展中心
所属专业机构	中国科学院北京基因组研究所(国家生物信息中心)
起始时间	2022-03-31
结束时间	2022-05-01

人员信息	
项目负责人姓名	██████████
项目负责人邮箱	██████████
联系人姓名	██████████
联系人邮箱	██████████

项目汇交计划	
文件系统编号	PLAN-322-1.docx
文件名称	汇交计划模板.docx
文件类型	初版计划
状态	审核未通过
提交时间	2022-04-11
操作	[修改] [下载] [邮寄地址]

**完成提交后，请等待管理员进行审核
(审核周期1-2工作日)**

**注意：管理员反馈审核意见前，您可随时
修改。请尽量避免多次修改**

更新项目信息

修改项目基本信息

修改

下载

邮寄地址

修改汇交计划



汇交计划审核情况反馈

- **如果审核通过：**可通过**[信息反馈]**查阅报送专业机构须知。您仍可通过**[更新项目信息]**修改除了“项目资助编号”之外的所有信息，包括更换联系人；

个人项目列表 (您有1信息未读)

系统编号: SDAS000322
项目编号:  项目名称:  项目负责人: 

① 完成&重新制定汇交计划 ② 上传盖双章版汇交计划 ③ 在线查询&出具证明

【状态说明】汇交计划审核通过：汇交计划报送专业机构须知详见 [\[信息反馈\]](#)。如仍需调整，请点击 [\[完成&重新制定汇交计划\]](#) 在线提交申请




【状态说明】汇交计划审核通过：汇交计划报送专业机构须知详见 [\[信息反馈\]](#)。如仍需调整，请点击 [\[完成&重新制定汇交计划\]](#) 在线提交申请



汇交计划盖章文件寄出

- 管理员完成汇交计划盖章后，会为您寄出汇交计划盖中心章文件，请注意查收。

个人项目列表（您有1信息未读）

系统编号: SDAS000322
项目编号:  项目名称:  项目负责人: 

① 完成&重新制定汇交计划 > ② 上传盖双章版汇交计划 > ③ 在线查询&出具证明

【状态说明】汇交计划已寄出：完成盖章后，请点击[\[上传盖双章版汇交计划\]](#)，上传完整版的盖双章汇交计划，更多资讯详见[\[信息反馈\]](#)

 信息反馈

【状态说明】汇交计划已寄出：完成盖章后，请点击[\[上传盖双章版汇交计划\]](#)，上传完整版的盖双章汇交计划，更多资讯详见[\[信息反馈\]](#)

汇交计划制定阶段**正式结束**，汇交
计划将**无法再进行修改**



项目汇交计划盖章扫描件（双章）归档

- 双章版汇交计划扫描件上传：当项目层的项目汇交计划完成所在单位盖章后，请点击“上传盖双章版汇交计划”完成文件上传。

个人项目列表 (您有1信息未读)

系统编号: SDAS000322
项目编号: [REDACTED]
项目名称: [REDACTED]
项目负责人: [REDACTED]

① 完成&重新制定汇交计划 → ② 上传盖双章版汇交计划 → ③ 在线查询&出具证明

【状态说明】汇交计划已寄出：完成盖章后，请点击[\[上传盖双章版汇交计划\]](#)，上传完整版的盖双章汇交计划，更多资讯详见[\[信息反馈\]](#)

 信息反馈



项目汇交计划盖章扫描件（双章）归档

- **双章版汇交计划扫描件上传：**当项目层的项目汇交计划完成所在单位盖章后，请点击“上传盖双章版汇交计划”完成文件上传。

主页 / SDAS000322 / 汇交计划上传

SDAS000322: 区域特异性神经干细胞的获取以及功能特性和应用的研究

● 注意事项

- 项目承担单位都需寄送两份给所对应专业机构，以便专业机构审核归档，具体寄送方式及细节请咨询对应专业机构
- 盖双章版汇交计划必须提交完整版，已方便您或是专业机构进行在线浏览

双章版汇交计划上传

将“盖中心章版汇交计划”，按专业机构要求完成项目负责人签字和项目承担单位盖章等流程，再将材料完整扫描版上传至系统归档

* 请上传最终版盖双方章的完整版汇交计划（仅接收pdf版）
请单击此处选择文件后，点击右侧按钮上传

上传

提交



双章版计划审核情况反馈

- 如果审核未通过：请根据[信息反馈]中的建议，点击[完成/重新制定汇交计划]或[查看/修改盖双章版汇交计划]任一按键进入浏览页面，点击“修改”进行完善。

个人项目列表 (您有2信息未读)

系统编号: SDAS000322
项目编号: [REDACTED]
项目名称: [REDACTED]
项目负责人: [REDACTED]

① 完成&重新制定汇交计划 ▶ ② 查看/修改盖双章版汇交计划 ▶ ③ 在线查询&出具证明

【状态说明】双章版计划审核未通过：请点击【查看/修改盖双章版汇交计划】进行修改，建议详见【信息反馈】

三 项目汇交计划

项目自产数据类型 组学原始测序下机数据；

文件系统编号	文件名称	文件类型	状态	提交时间	操作
PLAN-322-2.docx	项目汇交计划.docx	初版计划	审核通过	2022-04-11	下载 邮寄地址 申请撤回计划
PDOC-322-1.pdf	项目汇交计划-盖双章版.pdf	双章版计划	审核未通过	2022-04-11	修改 下载



完成提交后，请等待管理员进行审核
(审核周期1-2工作日)

注意：管理员反馈审核意见前，您可随时
修改。请尽量避免多次修改



双章版计划审核情况反馈

- **如果审核通过：**项目汇交计划的双方公章电子扫描件已归档，建议项目层在执行期内，督促数据递交方及时通过[数据提交入口 \(BIG-Sub\)](#) 进行科学数据提交。
- 盖章扫描件归档同时，为项目层**开启的实时查询数据提交情况功能**，请点击[\[在线查询&出具证明\]](#)实时查询和监督数据归档情况，以及申请出具证明。

个人项目列表 (您有2信息未读)

系统编号: SDAS000322	项目编号:	项目名称:	项目负责人:
① 完成&重新制定汇交计划	② 完成盖双章版计划上传	③ 在线查询&出具证明	【状态说明】您可通过 [在线查询&出具证明] 实时查询数据归档情况及申请出具证明；数据递交，请前往数据提交入口 (BIG-Sub)，更多资讯详见 [信息反馈]
			信息反馈



培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



国家基因组科学数据中心资源体系

The screenshot shows the homepage of the National Genomics Data Center (NGDC). At the top, there are logos for '国家基因组科学数据中心' (National Genomics Data Center) and 'NSTI'. A navigation bar includes '登录' (Login) and '语言 / Language'. Below the header, a banner states: '面向我国人口健康和社会可持续发展的重大战略需求, 国家基因组科学数据中心(NGDC), 作为国家生物信息中心(CNCB)的重要组成, 建立生命与健康大数据汇交存储、安全管理、开放共享与整合挖掘研究体系, 研发大数据前沿交叉与转化应用的新方法和新技术, 建成支撑我国生命科学发展、国际领先的基因组科学数据中心。' A search bar allows users to find bioprojects, biosamples, genes, proteins, tools, and databases. Below the search bar, a message says: 'CNCB-NGDC实现国家人类遗传资源信息一体化管理备份, 欢迎访问使用。' A row of icons provides quick access to: '提交' (Upload), '科学项目数据汇交' (Data Submission), '人类遗传资源信息管理备份' (Resource Management Backup), '序列搜索比对' (Sequence Search & Comparison), '新冠病毒信息库' (Coronavirus Information Database), and '文献库' (Literature Database). On the left, a sidebar titled '资源' (Resources) lists categories such as '原始数据', '基因组和变异', '基因表达', '非编码RNA', '表观基因组', '单细胞组学', '生物多样性和生物合成', '健康和疾病', '文献和教育', and '工具'. On the right, a section titled '热门资源' (Popular Resources) displays icons and names for various databases: BioCode (生物工具软件), BioProject (生物项目库), BioSample (生物样本库), GSA (组学原始数据归档库), GSA-Human (人类组学原始数据归档库), OMIX (多元数据归档库), GWH (基因组序列库), GVM (基因组变异库), Database Commons (生物数据库目录), GEN (基因表达数据库), MethBank (甲基化数据库), and BIT (生物信息在线分析平台).

- 原始组学数据归档库 (Genome Sequence Archive)
- 人类遗传资源管理库 (GSA-Human)
- 生命科学程序代码归档库 (BioCode)
- 基因组序列数据库 (Genome Warehouse)
- 基因组变异数据库 (Genome Variation Map)
- 甲基化数据库 (Methylation Bank)
- 多元数据归档库 (OMIX)



科学数据提交须知

请选择您要提交的数据类型

BIG Sub Quick Start Guide BIG Sub 使用说明

科技计划项目科学数据汇交

SDAS

科学项目数据汇交服务系统
提供汇交计划提交、实时数据归档情况查询、汇交证明出具等服务

数据提交入口BIG Sub

<https://bigd.big.ac.cn/gsub>

关于账号：

1. 建议谁的数据谁提交，即数据拥有者注册提交，方便各数据库反馈数据集处理进度和情况
2. 对于GSA-Human数据库，一定要具备高级职称人员或使用此类人员邮箱进行账号注册



科学数据提交须知

请选择您要提交的数据类型

BIG Sub Quick Start Guide BIG Sub 使用说明

BioProject
收集与共享生物学研究项目信息的资源库

BioSample
收集与共享实验相关的生物样本信息的资源库

Genome Sequence Archive
组学原始数据汇交、存储、管理与共享系统

GSA for Human
人类遗传资源组学原始数据库

Genome Warehouse
多物种基因组组装序列和注释数据归档库

GenBase
核酸和蛋白序列及注释数据归档库

OMIX
收集和共享表达谱、甲基化谱、蛋白组学、代谢组学等数据的资源库

Genome Variation Map
基因组序列变异信息汇交、管理与检索的资源库

BioCode
整合开源生物信息软件工具的数据库

Database Commons
生物数据库收集、分类、检索系统

SDAS
提供汇交计划提交、实时数据归档情况查询、汇交证明出具等服务

数据提交入口BIG Sub

<https://bigd.big.ac.cn/gsub>

关于deadline：

1. **专业机构的官方规定**，项目应在执行期结束后三个月内完成数据提交并通过科技部系统完成汇交凭证上传
2. **请不要等到结题再突击提交数据**，建议**随产生随提交**
3. 中心各数据库都可以支撑文章发表，可以**发文章前提交数据，将编号写在文章中**，既发文章又完成任务一举两得



数据库简要介绍

涉及数据库	主要数据类型	联系方式
<u>GSA</u>	非人类的组学测试下机数据	邮箱: gsa@big.ac.cn QQ群: 548170081
<u>GSA-Human</u>	人类的组学测试下机数据	说明文档: GSA使用说明 、 GSA-Human 使用说明
<u>OMIX</u>	基因组二级分析数据、代谢组测序数据、蛋白组测序及分析数据、蛋白结构数据、图像数据、芯片数据、临床信息等	邮箱: gsa@big.ac.cn QQ群: 516418530 说明文档: OMIX使用说明
<u>GWH</u>	全基因组、宏基因组等基因组组装序列和注释数据	邮箱: gwh@big.ac.cn QQ群: 183915274 , 541196594
<u>GVM</u>	基因组SNP变异、InDel变异、SNP芯片	邮箱: gvm@big.ac.cn QQ群: 468638108
<u>BioCode</u>	生物类软件工具和方法	邮箱: biocode@big.ac.cn QQ群: 537473603
<u>Database Commons</u>	生物类数据库	邮箱: databasecommons@big.ac.cn



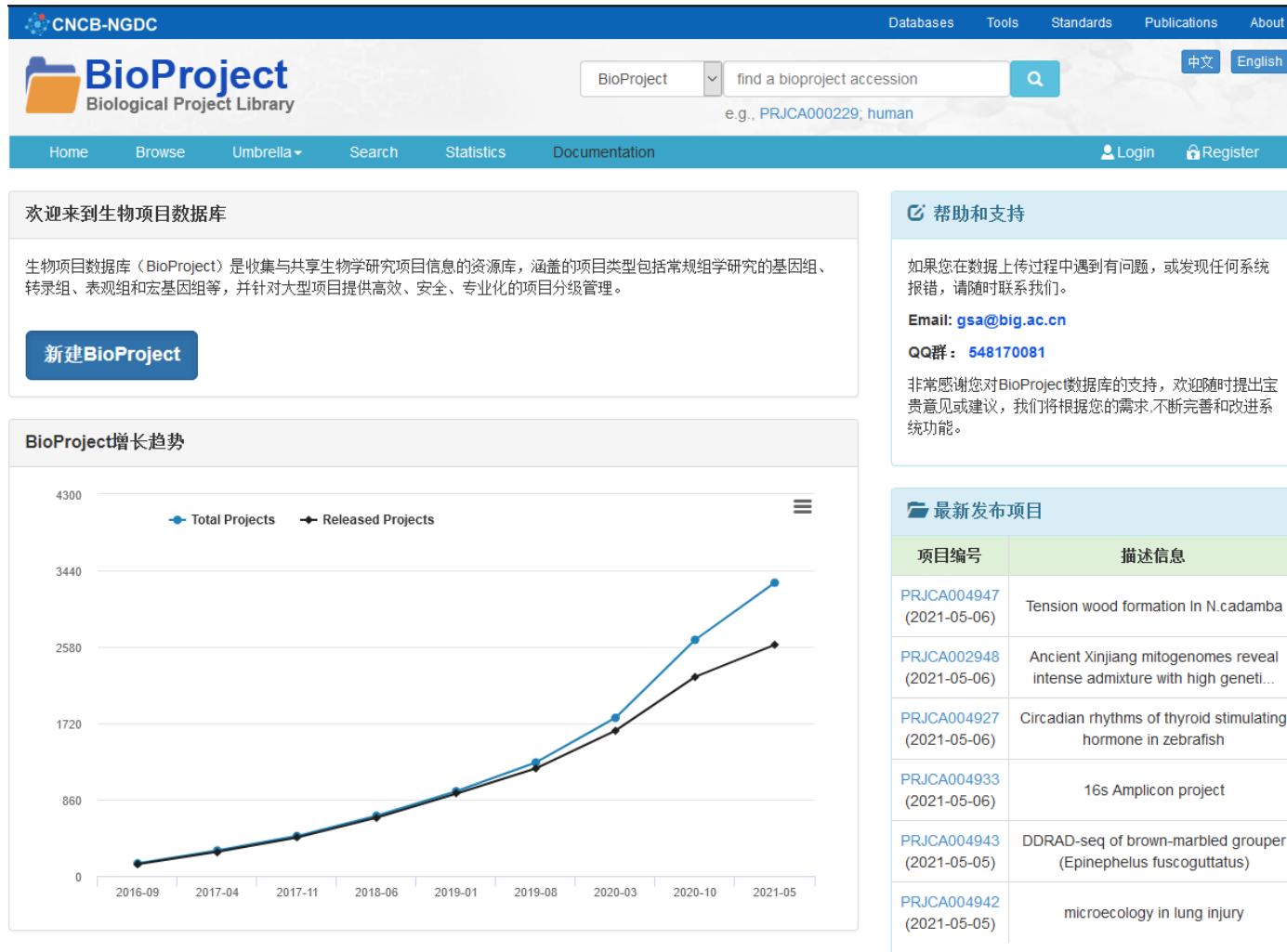
如何填写和调整国家重点专项项目资助号

- 只有正确关联了国家重点专项项目资助号信息的数据集，才会出现在最终的汇交证明中
 - GSA, GSA-Human, GWH, GVM和OMIX数据库都需要关联BioProject，请在BioProject创建时正确填写资助编号
 - Database Commons和BioCode，资助信息是收录在自己库中的，填写后可以随时进行修改

涉及数据库	联系方式
<u>BioProject</u>	邮箱: gsa@big.ac.cn QQ群: 548170081 说明文档: BioProject使用说明
<u>BioCode</u>	邮箱: biocode@big.ac.cn QQ群: 537473603
<u>Database Commons</u>	邮箱: databasecommons@big.ac.cn



BioProject数据库



BioProject主网页：

<https://ngdc.cncb.ac.cn/bioproject/>

说明文档

[BioProject使用说明](#)

问题反馈邮箱：

gsa@big.ac.cn



BioProject : 填写提交者信息

BIG Sub / BioProject / New BioProject

01 提交者信息 02 基本信息 03 项目类型 04 出版信息 05 概览 & 提交

提交者信息

* 名 Zhang	中间名 middle name	* 姓氏 Sisi
* 邮箱 zhangss@big.ac.cn	备用邮箱 secondary email	
* 单位 Beijing Institute of Genomics (BIG)	单位网址 http://www.big.ac.cn	* 部门 BIG Data Center
手机	传真	
* 街道 Beichen West Road	* 城市 Beijing	州/省
* 邮编 100101	* 国家/地区 China	
保存并进入下一项		



BioProject: 填写基本信息

BIG Sub / BioProject / New BioProject

01 提交者信息 02 基本信息 03 项目类型 04 出版信息 05 概览 & 提交

基本信息

发布日期

审核通过后即可发布
 指定日期发布
2024-09-23

(yyyy-mm-dd)

先导项目

—

项目

* 项目标题 * 涉及领域

* 项目说明

项目资金来源

* 机构 项目类别 * 项目批准号 项目名称

支持添加多个资助编号

GSA, GSA-Human, GWH, GVM和OMIX数据库都需要关联BioProject, 请在BioProject创建时正确填写资助编号

数据汇交用户, 需正确填写资助编号, 否则系统无法关联查询到数据

机构 (下拉菜单)

Ministry of Science and Technology of the People's Republic of China (MOST)	科技部项目
National Natural Science Foundation of China (NSFC)	基金委项目
Chinese Academy of Sciences (CAS)	中科院项目
Others	其他项目
No funding support	无资助



BioProject : 填写项目类型和样本范围

BIG Sub / BioProject / New BioProject

01 提交者信息 02 基本信息 03 项目类型 04 出版信息 05 预览 & 提交

项目类型

* 项目数据类型

- Whole genome sequencing
- Clone ends
- Epigenomics
- Exome
- Map
- Metagenome
- Phenotype or Genotype
- Random survey
- Targeted Locus (Loci)
- Transcriptome or Gene expression
- Variation
- Genome sequencing and assembly
- Raw sequence reads
- Genome sequencing
- Assembly
- Metagenomic assembly
- Targeted loci cultured
- Targeted loci environmental
- Single cell sequencing
- Other

* 样本范围

Monoisolate

样本范围选择提示 ●

Monoisolate：单个动物、培养细胞系、近交群体（或可能是从集合样本中产生单个基因型群体），注意此项适用范围较窄，不为优选项，请谨慎选择。

Multiisolate：同一物种的多基因型的个体样本集；

Multi-species：多物种样本集；

Environment：环境样本集，多用于宏基因组研究；

Synthetic：在实验室里制造/合成的样本集；

Single cell：单个细胞测序样本集；

Other：其他样本集/无法归入以上类别的特殊样本集。

保存并进入下一项

按需求填写项目数据类型信息，如
果可选项里没有，可点击“other”
自行添加



BioProject : 填写出版信息

BIG Sub / BioProject / New BioProject

01 提交者信息 02 基本信息 03 项目类型 04 出版信息 05 概览 & 提交

出版信息

PubMed ID OR DOI 

杂志名称

文章标题

发表年份 卷 期 起始页 结尾页

作者列表

+ Add another publication

保存并进入下一项



BioProject信息浏览&提交

BIG Sub / BioProject / New BioProject

01 提交者信息 02 基本信息 03 项目类型 04 出版信息 05 概览 & 提交

概况信息

提交者信息

提交者	Zhang Sisi zhangss@big.ac.cn
单位	Beijing Institute of Genomics (BIG)
部门	BIG Data Center
国家/地区	China
地址	Beichen West Road Beijing
邮编	100101

基本信息

项目标题	A phase I, multicenter study to assess								
涉及领域	Medical								
项目说明	By ate non-minor mutations in EGFR or HER2 DZD9008 was determined by oraladmisstration in patients with lung cancer Safety and tolerability								
发布日期	2024-08-02								
项目资金来源	<table border="1"><thead><tr><th>机构</th><th>项目类别</th><th>项目批准号</th><th>项目名称</th></tr></thead><tbody><tr><td>Ministry of Science and Technology of the People's Republic of China (MOST)</td><td></td><td>2017YFA0504100</td><td></td></tr></tbody></table>	机构	项目类别	项目批准号	项目名称	Ministry of Science and Technology of the People's Republic of China (MOST)		2017YFA0504100	
机构	项目类别	项目批准号	项目名称						
Ministry of Science and Technology of the People's Republic of China (MOST)		2017YFA0504100							

项目类型

填写资助编号信息会显示在此处

项目数据类型	Variation
样本范围	Monoisolate

出版信息

提交

请务必点击“**提交**”，提交后无需
等待审核，可以继续完成后续信息提交



培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



实时查询与监督数据归档情况

- 项目层可通过[\[在线查询&出具证明\]](#)进入实时数据归档情况的查询界面

个人项目列表 (您有2信息未读)

系统编号: SDAS000322
项目编号: [REDACTED]
项目名称: [REDACTED]
项目负责人: [REDACTED]

① 完成&重新制定汇交计划 → ② 完成盖双章版计划上传 → ③ 在线查询&出具证明

【状态说明】您可通过[\[在线查询&出具证明\]](#)实时查询数据归档情况及申请出具证明；数据递交，请前往数据提交入口（BIG-Sub），更多资讯详见[\[信息反馈\]](#)

[信息反馈](#)



实时查询与监督数据归档情况

- 项目层可通过[\[在线查询&出具证明\]](#)进入实时数据归档情况的查询界面

科技项目数据汇交服务系统

主页 / 个人中心 / 查询汇交数据

1 查询数据汇交情况 → 2 预览证明并签署协议 → 3 完成 →

查询数据汇交情况: [REDACTED]

接收数据资源清单

系统仅能展示“已关联项目资助编号”的数据集,您可点击“批量导出列表”按钮导出汇交情况表格,以便发送给数据提交者进行信息核对

显示: 5 项结果 搜索:

数据/文件编号	数据集名称	未归档数据集显示为粉红色	样本数/文件数	数据量	数据格式	共享方式	发布时间	递交人邮箱	所属数据库
[REDACTED]	[REDACTED]	1		fastq	公开访问	[REDACTED]	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	15		fastq	公开访问	[REDACTED]	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	33	116.07GB	fastq	公开访问	[REDACTED]	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	9	22.01GB	fastq	公开访问	[REDACTED]	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	1	45.60GB	fastq	公开访问	[REDACTED]	[REDACTED]	[REDACTED]	GSA

显示第 1 至 5 项结果, 共 17 项 上页 1 2 3 4 下页

保存并预览证明 返回完善数据 批量导出列表

请点击“批量导出列表”可导出并查看所有数据归档情况

系统仅展示正确关联“项目资助号”的数据集, 未正确填写编号的将不出现在列表中, 须知详见[“如何填写和调整国家重点专项项目资助号”](#)



如何看懂批量导出数据信息

- 以下为导出表格样例：表头含有常见问题、可能原因、建议问题处理器、处理方法和数据库联系方式等信息，未归档数据集显示为**红字+黄色高亮**，方便项目层进行反馈。

A	B	C	D	E	F	G	H	I	J	K
常见问题	可能原因	问题处理器	处理方法	联系方式						
1 数据集未在列表中	未关联资助编号	提交者	联系管理员开放BioProject编辑权限	邮箱gsa@big.ac.cn或QQ群：548170081						
2 有未归档数据集	数据未完成提交	提交者	联系管理员核查问题	邮箱gsa@big.ac.cn或QQ群：548170081						
3 有不属于项目的数据集	错误关联资助编号信息	提交者	联系管理员核查问题	邮箱gsa@big.ac.cn或QQ群：548170081						
4										
5										
6	列表仅能展示已关联项目资助编号数据集，如有黄红色标记的行，代表该数据集未归档									
7 数据/文件编号	状态	数据集名称	BioProject编号	样本/文件数	数据量	数据格式	共享方式	发布时间	递交人邮箱	所属数据库
8	未归档			1		fastq	公开访问		example@email.com	GSA
9	未归档			15		fastq	公开访问		example@email.com	GSA
10	已归档			33	116.07GB	fastq	公开访问		example@email.com	GSA
11	已归档			9	22.01GB	fastq	公开访问		example@email.com	GSA
12	已归档			1	45.60GB	fastq	公开访问		example@email.com	GSA

SDAS仅为项目层提供查询服务，数据处理问题，一定要请提交者联系数据库进行咨询和处理



在线浏览汇交证明

- 项目层组织数据递交责任方对数据提交完成情况进行核对，确保所需数据集都归档后，一定要确定核对无误后，再单击“保存并预览证明”进入下一步。

科技项目数据汇交服务系统

主页 / 个人中心 / 查询汇交数据

1 查询数据汇交情况 2 预览证明并签署协议 3 完成

查询数据汇交情况：[REDACTED]

接收数据资源清单

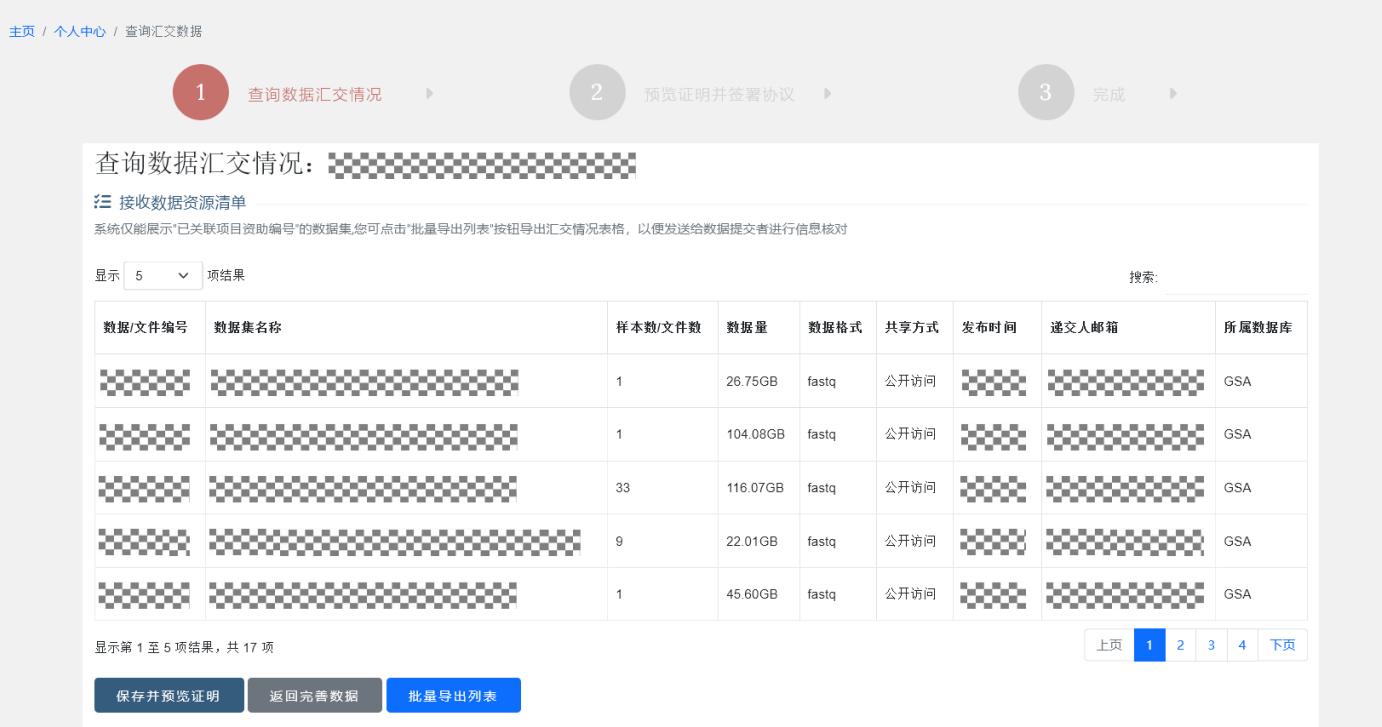
系统仅能展示“已关联项目资助编号”的数据集,您可点击“批量导出列表”按钮导出汇交情况表格,以便发送给数据提交者进行信息核对

显示 5 项结果 搜索:

数据/文件编号	数据集名称	样本数/文件数	数据量	数据格式	共享方式	发布时间	递交人邮箱	所属数据库
[REDACTED]	[REDACTED]	1	26.75GB	fastq	公开访问	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	1	104.08GB	fastq	公开访问	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	33	116.07GB	fastq	公开访问	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	9	22.01GB	fastq	公开访问	[REDACTED]	[REDACTED]	GSA
[REDACTED]	[REDACTED]	1	45.60GB	fastq	公开访问	[REDACTED]	[REDACTED]	GSA

显示第 1 至 5 项结果, 共 17 项 上页 1 2 3 4 下页

保存并预览证明 返回完善数据 批量导出列表



凭证与汇交计划不完全一致常见的，只要能自圆其说就可以



在线浏览汇交证明

- 预览终版证明，在线签署“承诺书”，并填写邮寄信息

科技项目数据汇交服务系统

主页 / 个人中心 / 汇交汇交证明

1. 查看数据汇交情况 > 2. 预览证明并签署协议 > 3. 完成 >

出具汇交证明:

项目汇交证明预览

No. NGDC-2022-0266

科技计划项目数据汇交证明

汇交证明生成协议

科学数据汇交证明出具承诺书

1. 本项目保证所有汇交的数据已归入国家基因组科学数据中心（下文简称“中心”）平台管理，不会随意修改和删除汇交的科学数据。
2. 本项目保证所有汇交的数据真实、准确、完整、有效，无学术不端和学术失范行为，并承担因数据不真实等问题带来的的一切后果和法律责任。
3. 本项目保证以真实身份签署证明，确定所提交本作为最终版物收证明。

我已认真核对经此汇交证明内容，确定申请生成该文件
我已阅读了汇交证明生成协议，同意并遵守所有条款

盖章件邮寄地址

使用汇交计划的邮寄信息

+ 收件人姓名 * 收件地址
例如：北京市朝阳区光华东街1号院104号楼

* 手机号码

备注

保存并预览 > 返回上一步



在线浏览汇交证明

- 预览终版证明，在线签署“承诺书”，并填写邮寄信息

科技项目数据汇交服务系统

主页 / 个人中心 / 预览汇交证明

1. 预览数据汇交情况 2. 预览证明并签署协议 3. 完成

出具汇交证明：███████████

项目汇交证明预览

No. NGDC-2022-0266

科技计划项目数据汇交证明

汇文证明生成协议

科学数据汇交证明出具承诺书

1. 本项目保证所有数据已纳入国家基因组科学数据中心（下文简称“中心”）平台管理，不会擅自修改和删除汇交的科学数据；
2. 本项目保证所有汇交的数据真实、准确、完整、有效，无学术不端和学术失范行为，并承担因数据不真实等问题带来的后果和法律责任；
3. 本项目确定已认真核对以上验收证明，确定现有版本作为最终版验收证明。

□ 我已认真核对终版汇交证明内容，确定申请生成盖章文件
□ 我已经阅读了汇交证明生成协议，同意并遵守所有条款

盖章件邮寄地址

使用汇交计划的邮寄信息

* 收件人姓名 * 手机号码 * 收件地址

备注

例如：北京市朝阳区北辰西路1号院104号楼

保存并提交 返回上一步

汇交证明生成协议

科学数据汇交证明出具承诺书

1. 本项目保证所有数据已纳入国家基因组科学数据中心（下文简称“中心”）平台管理，不会擅自修改和删除汇交的科学数据；
2. 本项目保证所有汇交的数据真实、准确、完整、有效，无学术不端和学术失范行为，并承担因数据不真实等问题带来的后果和法律责任；
3. 本项目确定已认真核对以上验收证明，确定现有版本作为最终版验收证明。

我已认真核对终版汇交证明内容，确定申请生成盖章文件
 我已经阅读了汇交证明生成协议，同意并遵守所有条款

盖章件邮寄地址

使用汇交计划的邮寄信息

* 收件人姓名 * 手机号码 * 收件地址

备注

例如：北京市朝阳区北辰西路1号院104号楼



汇交凭证审核情况反馈

- 如果审核未通过：请根据[信息反馈]中的建议，点击[在线查询&出具证明]进入浏览页面，通过[更新/浏览数据集]重新出具汇交证明。

The screenshot shows a user interface for managing project reviews. At the top, it displays the system number SDAS000322, a redacted project ID, a redacted project name, and a redacted responsible person. Below this, there is a navigation bar with three steps: 1. 完成&重新制定汇交计划, 2. 完成盖双章版计划上传, 3. 在线查询&出具证明. A red box highlights the third step. To the right of the third step is a blue button labeled '信息反馈'. Below the navigation bar, a message box contains the text: 【状态说明】证明审核未通过：可点击 [在线查询&出具证明] 进入浏览页面，点击[更新/浏览数据集]重新出具汇交证明，更多资讯详见 [信息反馈]. A red arrow points from this message box down to another identical message box at the bottom of the page, which is also enclosed in a red border.

个人项目列表

系统编号: SDAS000322

项目编号: [REDACTED]

项目名称: [REDACTED]

项目负责人: [REDACTED]

① 完成&重新制定汇交计划 > ② 完成盖双章版计划上传 > ③ 在线查询&出具证明

【状态说明】证明审核未通过：可点击 [在线查询&出具证明] 进入浏览页面，点击[更新/浏览数据集]重新出具汇交证明，更多资讯详见 [信息反馈]

信息反馈

【状态说明】证明审核未通过：可点击 [在线查询&出具证明] 进入浏览页面，点击[更新/浏览数据集]重新出具汇交证明，更多资讯详见 [信息反馈]



汇交凭证审核情况反馈

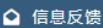
- 如果审核未通过：请根据[信息反馈]中的建议，点击[在线查询&出具证明]进入浏览页面，通过[更新/浏览数据集]重新出具汇交证明。

个人项目列表

系统编号: SDAS000322
项目编号: [REDACTED]
项目名称: [REDACTED]
项目负责人: [REDACTED]

① 完成&重新制定汇交计划 > ② 完成盖双章版计划上传 > ③ 在线查询&出具证明

【状态说明】证明审核未通过：可点击 [在线查询&出具证明] 进入浏览页面，点击[更新/浏览数据集]重新出具汇交证明，更多资讯详见 [信息反馈]

 信息反馈



项目汇交证明

证明编号	文件名称	文件类型	状态	提交时间	操作
NGDC-2022-0266	[REDACTED]	汇交证明	审核未通过	2022-04-11	 

 点击“更新/浏览数据集”更新证明



汇交凭证审核情况反馈

- **如果审核通过：**开始中心内盖章审批流程，汇交凭证盖章件将于1-2个工作日内寄出。

个人项目列表 (您有2信息未读)

系统编号: SDAS000322
项目编号: [REDACTED]
项目名称: [REDACTED]
项目负责人: [REDACTED]

① 完成&重新制定汇交计划 ▶ ② 完成盖双章版计划上传 ▶ ③ 汇交证明已出具

【状态说明】汇交证明已出具：盖章版汇交证明将于1-2个工作日内寄出。

✉ 2 信息反馈

【状态说明】汇交证明已出具：盖章版汇交证明将于1-2个工作日内寄出。



凭证盖章件寄出

- 凭证寄出后，状态改为“**汇交证明已寄出**”您即可通过系统浏览与下载“**凭证盖章扫描件**”。预祝项目验收顺利。

个人项目列表 (您有3信息未读)

系统编号: SDAS000322
项目编号: [REDACTED]
项目名称: [REDACTED]
项目负责人: [REDACTED]

① 完成&重新制定汇交计划 ▶ ② 完成盖双章版计划上传 ▶ ③ 汇交证明已出具

【状态说明】汇交证明已寄出：汇交证明已正式出具并寄出，预祝您们项目验收顺利

3 信息反馈



项目汇交计划

项目自产数据类型	文件名称	文件类型	状态	提交时间	操作
PLAN-25-1	[REDACTED]	初版计划	审核通过	2022-04-12	下载 邮寄地址
PDOC-25-1	[REDACTED]	双章版计划	审核通过	2022-04-12	下载

项目汇交证明

证明编号	文件名称	文件类型	状态	提交时间	操作
NGDC-2021-0056	[REDACTED]	汇交证明	已归档	2022-04-15	下载盖章证明 邮寄地址

汇交凭证一旦出具不予撤回，请您务必谨慎



培训主要内容

- 数据汇交政策背景
- SDAS系统功能概述
- 汇交计划制定与审核
- 科学数据提交须知
- 进度查询与凭证出具
- 常见问题与文档下载



帮助信息与文档下载

The screenshot shows the 'Help' section of the SDAS system. On the left, there's a sidebar with links for 'Common Questions', 'Project Submission Plan Preparation' (which is highlighted with a red border), 'Data Submission', 'Proof of Submission', 'Document Download', and 'Contact Us'. The main content area is titled 'Project Submission Plan Preparation' and lists seven frequently asked questions:

- 什么属于科学数据?
- 请问贵中心数据汇交的具体步骤?
- SDAS系统的主要功能?
- SDAS系统必须使用项目负责人注册吗? 还是项目联系人也可以?
- 如何更换项目基本信息中的联系人信息?
- 如何制定与修改汇交计划?
- 如何提交项目汇交计划的双方公章电子扫描件?

At the bottom, it says '显示第1至7个常见问题, 共7个' (Showing the first 7 common questions, a total of 7) and has navigation buttons for '上页' (Previous page), '1' (Current page), and '下页' (Next page).

<https://ngdc.cncb.ac.cn/sdas/help>

常见问题

- 项目汇交计划制定



帮助信息与文档下载

The screenshot shows a web page with a blue header bar. On the left, there are two logos: 'CNCB' and 'NGDC'. The main title is '科技项目数据汇交服务系统'. Below the title, there are several navigation links: '主页', '提交', '帮助' (highlighted in red), '注册', and '登录'. A sidebar on the left contains links for '常见问题', '项目汇交计划制定' (highlighted in red), '科学数据递交' (highlighted in red), '项目汇交证明出具', '文档下载', and '联系我们'. The main content area is titled '项目汇交计划制定' and '科学数据递交'. It lists four common questions under '科学数据递交':

- 如何提交数据, 应该递交到哪个数据库?
- 我作为项目层, 下属的课题/子课题都需要提交数据。我们应该使用什么账户, 或谁的账号递交数据?
- 我什么时候递交数据合适? 有没有deadline?
- 如何填写和调整国家重点专项项目资助号?

At the bottom of this section, it says '显示第 1 至 4 个常见问题, 共 4 个' and has buttons for '上页', '1' (highlighted in red), and '下页'.

<https://ngdc.cncb.ac.cn/sdas/help>

常见问题

- 项目汇交计划制定
- 科学数据递交



帮助信息与文档下载

The screenshot shows the 'Help' section of the 'Science Project Data Submission Service System'. The left sidebar includes links for 'Common Issues', 'Project Submission Plan Preparation' (highlighted with a red box), 'Scientific Data Submission', 'Project Submission Proof Issuance' (also highlighted with a red box), 'Document Download', and 'Contact Us'. The main content area is titled 'Project Submission Plan Preparation' and contains a 'Common Issues' section. This section is organized into a tree structure under 'Scientific Data Submission'. The first node is 'Project Submission Proof Issuance', which has two expanded items: 'How to determine if affiliated sub-topics have been submitted as a project layer?' and 'How to apply for a completion certificate online?'. A red box highlights the 'Project Submission Proof Issuance' node and its first two children. Navigation buttons at the bottom right of the content area include 'Previous', '1' (selected), and 'Next'.

<https://ngdc.cncb.ac.cn/sdas/help>

常见问题

- 项目汇交计划制定
- 科学数据递交
- 项目汇交凭证出具



帮助信息与文档下载

The screenshot shows the SDAS help page. At the top, there are navigation links for Databases, Tools, Standards, Publications, and About. Below that is a main menu with Home, Submit, Help, Register, and Login options. On the left, a sidebar includes links for Common Questions, Project Submission Plan, Scientific Data Submission, Project Submission Proof-of-Concept, Document Download (which is highlighted with a red box), and Contact Us. The main content area is titled 'Project Submission Plan' and contains sections for 'Scientific Data Submission' and 'Project Submission Proof-of-Concept'. Under 'Project Submission Proof-of-Concept', there is a 'Document Download' section which is also highlighted with a red box. This section lists 'Submission Plan Reference Template' (with download links for National Key Research Program and Brain Science and Intelligence Research Major Projects), 'SDAS System Usage Guide' (with download link), and 'National Genome Science Data Center Data Submission Guide' (with download link). There are also 'Display First' and 'Display Last' buttons.

<https://ngdc.cncb.ac.cn/sdas/help>

常见问题

- 项目汇交计划制定
- 科学数据递交
- 项目汇交凭证出具

文档下载

- 汇交计划参考模板
- 系统操作指南





THANKS

欢迎使用SDAS服务系统
<https://ngdc.cncb.ac.cn/sdas/>

工作邮箱: data@big.ac.cn



中国科学院北京基因组研究所（国家生物信息中心）
BEIJING INSTITUTE OF GENOMICS CHINESE ACADEMY OF SCIENCES / CHINA NATIONAL CENTER FOR BIOINFORMATION



国家基因组科学数据中心
National Genomics Data Center

生命科学程序代码归档库

BioCode

国家生物信息中心
China National Center for Bioinformation

工具提交者需提前准备的信息内容

- 软件工具基本描述性元信息
- 作者信息、相关文章信息
- 软件版本及源代码文件
- 软件工具说明文档
- 设定好软件信息公开时间
- 项目资助号信息



BIG Sub
BIG Submission Portal

Home Documentation

请选择您要提交的数据类型

BIG Sub Quick Start Guide BIG Sub 使用说明

生物数据递交系统(BIG Submission, BIG Sub)是国家基因组科学数据中心生物数据统一汇交入口,为用户提供一站式数据递交服务。

BioProject
收集与共享生物学研究项目信息的资源库

BioSample
收集与共享实验相关的生物样本信息的资源库

Genome Sequence Archive
组学原始数据汇交、存储、管理与共享系统

GSA for Human
人类遗传资源组学原始数据库

Genome Warehouse
多物种基因组组装序列和注释数据归档库

GenBase
核酸和蛋白序列及注释数据归档库

OMIX
收集和共享表达谱、甲基化谱、蛋白组学、代谢组学等数据的资源库

Genome Variation Map
基因组序列变异信息汇交、管理与检索的资源库

BioCode
整合开源生物信息软件工具的数据库

Database Commons
生物数据库收集、分类、检索系统



BioCode Categories Search Statistics

BioCode

Archive Bioinformatics Codes for Open Source Pro

Basic Statistics

Tools:	7210
Users:	24330
Organizations:	7074
Total Download Count:	242850

Most Popular Tools

- DeeReCT-TSS_release** DeeReCT-TSS: A no...
Tool Type: Application
Technologies: Python3 • Download Count: 59
- ILnc** Prioritizing Long Non-coding RNAs for Pan-can...
Tool Type: Application
Technologies: R • Download Count: 3863
- WBSA** Web Service for Bisulfite Sequencing Data A...
Categories DNA methylation • Tool Type: Pipeline & Protocol
Technologies: Perl, R • Download Count: 0
- GIREMI** Identify RNA editing sites
Categories RNA editing • Tool Type: Application
Technologies: C, Perl, R • Download Count: 0
- CandiHap** A haplotype analysis toolkit for natural v...
Categories Variant effect prediction • Tool Type: Toolkit
Technologies: Perl, Python2, R • Download Count: 27484
- Dr. Sim** Dr.Sim is a general learning-based framew...
Categories Drug repositioning • Tool Type: Application
Technologies: Python2, Python3 • Download Count: 648
- SNP Calling & ChIP-Seq analysis pipeline...**
Categories SNP detection, De novo ChIP-seq analysis • Tool Type: Pipeline & Protocol
Technologies: BASH • Download Count: 0

Latest Updated Tools

- PanGraphViewer**: v1.02
Date: September 17, 2022
- VP-Detector** 三维颗粒定位和分类软件: VPD1.0
Date: September 6, 2022
- 三维重构体图像去噪软件: SCNet1.0**
Date: September 6, 2022
- 二维冷冻电镜图像去噪软件: NT2C1.0**
Date: September 6, 2022
- 高通量原位结构解析软件包: isSPA1.0**
Date: September 6, 2022
- 分块重构软件包: Ver1.0**
Date: September 6, 2022
- NetBCE: 1.0**
Date: August 17, 2022
- stAPAMiner: v1.0**
Date: August 3, 2022
- DogVC: v1.0**
Date: August 2, 2022

登录 (Sign In)
注册 (Register)

<https://ngdc.cncb.ac.cn/gsub>

<https://ngdc.cncb.ac.cn/biocode>

提交按钮

CNCB-NGDC

Databases Tools Standards Publications About

BioCode Categories Search Statistics

Submit Tool Hao Lili

BioCode

Archive Bioinformatics Codes for Open Source Projects

Search

Basic Statistics	
Tools:	6997
Users:	24080
Organizations:	6818
Total Download Count:	13653

Most Popular Tools	
WBSA	Web Service for Bisulfite Sequencing Data... Categories DNA methylation • Tool Type: Pipeline & Protocol Technologies: Perl, R • Download Count: 0
GIREMI	Identify RNA editing sites Categories RNA editing • Tool Type: Application Technologies: C, Perl, R • Download Count: 0
CandiHap	A haplotype analysis toolkit for natural... Categories Variant effect prediction • Tool Type: Toolkit Technologies: Perl, Python2, R • Download Count: 3424

Latest Updated Tools	
VirusMuT : 1.0	Date: March 17, 2021
CVTree : 3.0.1	Date: March 5, 2021
Asymptomatic transmission of SARS-CoV-2 : 1.0	Date: February 1, 2021
SARS-CoV2-iSNV : 1.0	Date: February 1, 2021

信息提交界面

受控可选项

软件名称

简要描述

网站URL

公开时间

CNCB-NGDC

BioCode Categories Search Statistics

Name
Tool name

Description
Short description

Website
<http://example.com>

Public Available Date
<http://example.com>

Technologies

Platforms

User Interfaces

Tool Type
Application

Input Data Types

Categories

Save

Introduction
Formats **B I U Ix A A**

Publications
No publications attached

Fundings
资助项目号

Credits
Email
haolili@big...
Full Name
Hao Lili
Roles
Investigator
Organization
Beijing Institutes of Genomics
Department
Key Laboratory of Genome Sciences and Information
Country/Region
China

Submit Tool Hao Lili

相关文章

作者信息

受控可选项内容

Technologies

- BASH
- C
- C++
- GPU
- Hadoop

开发语言

Platforms

- Linux/Unix
- MAC OS X
- Windows

User Interfaces

- Desktop GUI
- Terminal Command Line
- Webpage

Tool Type

- Application
- Application
- Framework
- Infrastructure
- Pipeline & Protocol
- Toolkit

工具类型

Input Data Types

- BAM
- FASTA
- FASTQ
- SAM
- VCF

Categories

- High-throughput sequencing
- Base calling
- Read quality control
- Error correction
- Duplicate read removal
- Adapter trimming



CNCB-NGDC

BioCode Categories Search Statistics

Databases Tools Standards Publications About

Submit Tool Hao Lili

BioCode

Archive Bioinformatics Codes for Open Source Projects

查看已提交软件信息

Submitted Tools

Logout

This screenshot shows the main landing page of the BioCode platform. At the top, there's a navigation bar with links for Databases, Tools, Standards, Publications, and About. Below that is a secondary navigation bar with BioCode, Categories, Search, and Statistics. On the right side, there are buttons for Submit Tool and a user profile for 'Hao Lili'. A large orange arrow points down from a callout bubble labeled '查看已提交软件信息' (View Submitted Software Information) to the 'Submitted Tools' button.

CNCB-NGDC

BioCode Categories Search Statistics

Databases Tools Standards Publications About

Submit Tool Hao Lili

Name	Description	Updated At	Operations
RED-ML	RED-ML is a software tool to do genome-wide RNA editing detection (RED) based on RNA-seq data. All source codes and executables are located in the "bin" directory. The tool can be run on a Linux platform and the main program is red_ML.pl.	April 7, 2018	Edit

This screenshot shows a list of submitted tools. It includes columns for Name, Description, Updated At, and Operations. The 'Edit' button in the Operations column for the 'RED-ML' entry is highlighted with a red box and a callout bubble containing the text '点击“Edit”编辑已保存软件信息' (Click "Edit" to edit saved software information).

RED-ML RED-ML is a software tool to do genome-wide RNA editing detection (RED) based on RNA-seq...

[Home](#) [Manual](#) [Downloads](#) [Statistics](#)



已保存基本信息条目
皆可编辑修改保存

CNCB-NGDC

BioCode Categories Search Statistics

Submit Tool Hao Lili

RED-ML

RED-ML is a software tool to do genome-wide RNA editing detection (RED) based on RNA-se...

Manual Home Downloads Statistics

[Edit](#)

Manual

Parameters

```
--rnabam [STR] the sorted BAM file obtained from RNA-seq to detect RNA editing sites.  
--reference [STR] the fasta file containing the reference genome, e.g., hg19.fa.  
--dbSNP [STR] the SNP database file, e.g., dbSNP138.  
--simpleRepeat [STR] genome-wide simple repeat annotation, should be in BED format.  
--alu [STR] genome-wide Alu repeat annotation, should be in BED format.  
--snplist [STR] a tab-delimited file listing known SNPs, with the first two columns being chromosome and position of each SNP [optional].  
--outdir [STR] the directory of output.  
-p [NUM] the detection threshold, a number between 0 and 1 [default 0.5];  
-help [STR] show this help information!
```

Examples

We have provided a simple example to test the installation of RED-ML. Under the "example" directory, run:

```
perl ..bin/red_ML.pl --rnabam example.rna.bam --reference /usr/hg19.fa --dbSNP example.dbSNP.vcf --simpleRepeat example.simpleRepeat.bed --alu example.alu.bed --outdir ./test/
```

It should finish running in ~2 minutes with three output files (RNA_editing.sites.txt, variation.sites.feature.txt and mut.txt.gz). Here is another example of using RED-ML:

```
perl red_ML.pl --rnabam in.bam --reference hg19.fa --dbSNP dbSNP138.vcf --simpleRepeat hg19_simpleRepeat.reg.bed --alu hg19.alu.bed --snplist snp.list --outdir outdir
```

Requirements

RED-ML requires the following data files at the time of public release:

The reference genome (hg19), downloaded from: <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/chromosomes>.
dbSNP138, downloaded from: <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database>.
simpleRepeat, downloaded from: <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database>, and then do:
awk '{print \$2"\t"\$3"\t"\$4}' simpleRepeat.txt > simpleRepeat.bed
bedtools merge -i simpleRepeat.bed > simpleRepeat.merge.bed
Alu, downloaded from: <http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database>, and do:
grep Alu rmsk.txt | awk '{print \$6"\t"\$7"\t"\$8}' > hg19.alu.bed

We have also provided the simpleRepeat and Alu files under the "database" directory for the user's convenience.

Optional Steps

SNP calling

If you have matching DNA-seq data or aligned DNA BAM files, we strongly recommend to take advantage of them. You could call SNPs by GATK (haplotypecaller) or SOAPsnp and modify the format of the resulting file (such as vcf) to fit the format required by --snplist.

Alignment

Although RED-ML can accept BAM files produced by different alignment tools, the current version has been optimized for BWA and TopHat2 during the construction of our ML model, and we find that the choice of alignment tools and the associated parameters could have a large impact on RED. To help users with proper alignment strategies, we recommend the following steps:

- When reads are aligned by BWA (preferred), one should first build a new reference which combines reference genome (hg19) and exonic sequences surrounding all known splice junctions, and the detail method is the same as in Ramaswami et al. (Nature Methods 2012) and Wang et al. (GigaScience 2016). SAMtools can be used to sort the alignment file and remove the PCR duplicate reads.
- When TopHat2 is chosen, the cleaned reads can be mapped to the reference genome (hg19) directly with default parameters. Picard should be used to sort the alignment and to remove duplicate reads induced by PCR, and base quality score recalibration can be carried out by GATK.

Outputs

When the program finishes running, three files will be created in the output directory. RNA_editing.sites.txt lists all detected RNA editing sites that pass the detection threshold p; variation.sites.feature.txt lists all variant sites with associated feature values; mut.txt.gz contains all variant sites with pileup information.

点击‘Manual’可填写和修改软件说明文档

RED-ML

RED-ML is a software tool to do genome-wide RNA editing detection (RED) based on RNA-se...

[Home](#) [Manual](#)[Downloads](#)[Statistics](#)

Releases

1.0

April 7, 2018

Files Download Count: 361

- RED-ML-master.zip

[Edit](#)[+ New Release](#)

增加软件新版本信息

点击‘Downloads’可填写软件版本信息并上传软件元码

RED-ML

RED-ML is a software tool to do genome-wide RNA editing detection (RED) based on RNA-se...

Version: 1.0

Note:

Formats: B I U Zx A A

Upload: 浏览... 未选择文件。

Save

修改现有软件版本信息

File Name	File Size	Delete
RED-ML-master.zip	14.2 MB	Delete

RED-ML

RED-ML is a software tool to do genome-wide RNA editing detection (RED) based on RNA-se...

[Home](#)[Manual](#)[Downloads](#)[Statistics](#)

Introduction

RED-ML: RNA Editing Detection basedon Machine learning (pronounced as "red ML"). The input to RED-ML can be as simple as a single BAM file, while it can also take advantage of matched genomic variant information when available. The output not only contains detected RNA editing sites, but also a confidence score to facilitate downstream filtering. The developers have carefully designed validation experiments and performed extensive comparison and analysis to show the efficiency and effectiveness of RED-ML under different conditions, and it can accurately detect novel RNA editing sites without relying on curated RNA editing databases. This tool is also freely available via GitHub <<https://github.com/BGIRED/REDML>>.

Publications

- Red MI: A Novel, Effective Rna Editing Detection Method Based On Machine Learning. [Cite this](#)

Xiong H, Liu D, Li Q, Lei M, Xu L, Wu L, Wang Z, Ren S, Li W, Xia M, Lu L, Lu H, Hou Y, Zhu S, Liu X, Sun Y, Wang J, Yang H, Wu K, Xu X, Lee LJ, 2017 May 1 - *Giga Science*

Credits

- Leo J Lee ljlee@psi.toronto.edu Department of Electrical and Computer Engineering, University of Toronto, Canada
- Xun Xu xuxun@genomics.cn China National GeneBank, BGI-Shenzhen, China

Community Ratings

Usability	Efficiency	Reliability	Rated By
			1 users
			hao***i@big.ac.cn (May 7, 2021)

Summary	
Accession	BT007072
Tool Type	Application
Category	RNA editing
Platforms	Linux/Unix
Technologies	Perl
User Interface	Terminal Command Line
Input Data	BAM, FASTA
Latest Release	1.0 (April 7, 2018)
Download Count	361
Country/Region	Canada
Submitted By	Hao Lili

关联项目资助号

Fundings
This project is supported by the Shenzhen Peacock Plan (NO. KQTD20150330171505310).



THANKS

欢迎使用BioCode系统
<https://ngdc.cncb.ac.cn/biocode/>

工作邮箱：biocode@big.ac.cn

QQ群：537473603