中国人类遗传资源
个性信息描述规范编写规则

Compiling Formula for Description Criteria

for the particular Information of Chinese Genetic Resources

（讨论稿）

www.egene.org.cn

中国人类遗传资源平台项目组

二〇〇六年六月

相关信息

版权所有：中国人类遗传资源平台

电话：86-10-62179059

邮件： nicgr@egene.org.cn

主题：中国人类遗传资源信息描述规范

名称：人类遗传资源个性描述规范编写规则

编号：NIPCGR- DCI-04

www.egene.org.cn
www.egene.org.cn

目　次

前　　言 1
1 范围 2
2 术语和定义 3
3 规范性引用文件 5
4 编写方法 6
4.1 扩展要求 6
4.2 专题性要求 6
4.3 字段描述要求 7
4.4 值域要求 7
5 格式 8
5.1 标题 8
5.2 前言 8
5.3 正文 8
5.3.1 范围 8
5.3.2 规范性引用文件 8
5.3.3 定义 8
5.3.4 原则 8
5.3.5 方法 8
5.3.6 个性规范描述表 8
5.3.7 个性描述规范简表 9
5.3.8 个性描述规范字段说明 9
5.3.9 附录 1：字段赋值代码表 9
5.3.10 附录 2：描述规范录入示例 9
6 规范性附录：人类遗传资源特性信息描述规范及编码标准简表 10

www.egene.org.cn
www.egene.org.cn

前 言

中国人类遗传资源平台建设是一个以人类遗传资源信息共享为中心，提供技术支持和应用研究网络共享平台信息系统，其核心在于建立人类遗传资源数据库群。数据库群的建立必须要有严格的描述规范，才能保证数据的可用性、有效性和一致性，满足相关管理部门的业务要求。

为了指导各课题参加单位能更好地完成个性描述规范的编写，特制定《中国人类遗传资源平台个性描述规范编写规则》，供各单位参考。

本规范为第一次制定。

www.egene.org.cn

1 范围

本规范规定了如何在人类遗传资源特性信息描述规范及编码标准的基础上制定各个专题领域的人类遗传资源个性描述规范的一般方法。

www.egene.org.cn
www.egene.org.cn

2 术语和定义

2.1 人类遗传资源特性信息描述规范及编码标准

指在人类遗传资源共享平台中制定的用于描述各种人类遗传资源样本的规范，直观地来看，它是一系列的人类遗传资源数据库的数据字典。

人类遗传资源特性信息描述规范及编码标准是自然科技资源共享平台中对自然科技资源共性及身份描述的统一规范，包括各类资源共性的护照信息、实物信息、基本信息、特征信息、采集信息、关联信息、图像信息和共享信息等资源共性信息描述规范。

采用人类遗传资源特性信息描述规范及编码标准标准化整理后的数据进入自然科技资源 e-平台后，可向社会发布国家自然科技资源的种类、数量、分布及获取方式等基本信息，为资源的社会共享提供有效服务。

2.2 专用标准

标准术语，指由一个标准泛化形成另一个标准。如，A 标准泛化形成 B 标准，则称 B 标准为 A 标准的专用标准（Profile）。

2.3 人类遗传资源个性描述规范

人类遗传资源个性描述规范是人类遗传资源共享平台中特定资源所有特有的描述规范。该规范只适用于特定主题的数据库。

2.4 人类遗传资源描述规范系列

人类遗传资源共享信息系统中的人类遗传资源特性信息描述规范及编码标准和多个个性描述规范共同组成了人类遗传资源描述规范系列。

2.5 课题数据库

课题数据库指人类遗传资源共享信息系统项目参加单位根据个性描述规范制作的数据库。

www.egene.org.cn
www.egene.org.cn

2.6 样本数据库

指依据人类遗传资源特性信息描述规范及编码标准，从课题数据库中抽取出来的数据库。

2.7 专题数据库

课题数据库中提取样本数据库之后，形成的专题数据库称为专题数据库。

2.8 共性数据库

共性数据库指按照人类遗传资源共性描述规范，从主体数据库中提取的数据库。

www.egene.org.cn
www.egene.org.cn

3 规范性引用文件

下列文件中的条款通过本规范的引用而成为本规范的条款。凡是注明日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本规范，然而，鼓励根据本规范达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本规范。

GB/T1.1-2000　　标准的结构和编写规则

www.egene.org.cn

4 编写方法

4.1 扩展要求

个性描述规范必须是在《人类遗传资源特性信息描述规范及编码标准》基础上进行编写，个性数据库字段须是人类遗传资源特性信息描述规范及编码标准均不包含的字段。描述规范系列于系列数据库之间的关系见下图。

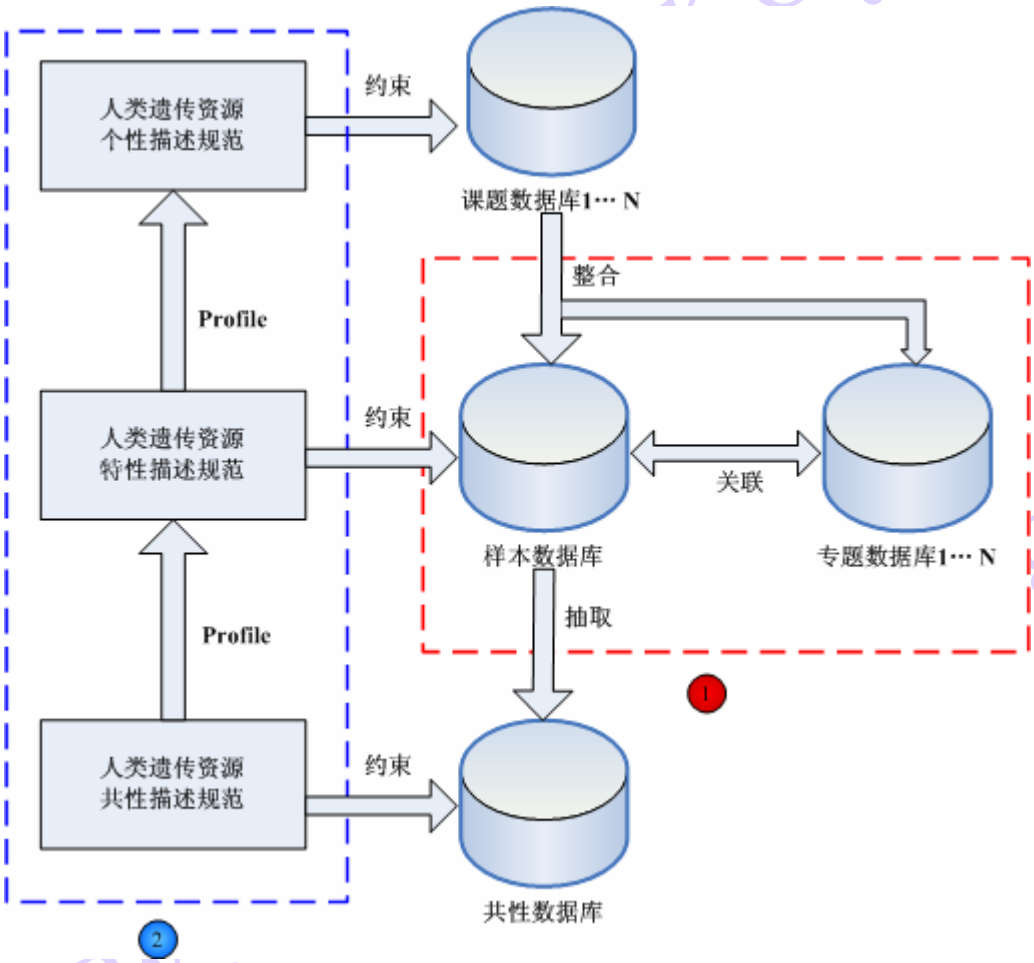

图 描述规范系列与系列数据库之间的关系

4.2 专题性要求

描述规范提出的数据库字段应充分体现个性研究领域的特征，数据库字段能够反应描述领域的特点和要求，同时考虑字段对资源查询者的实用性和易于理解性。

4.3 字段描述要求

需要描述的信息需提出准确的字段名称，方便资源查询者能够理解字段名称所描述的个性描述信息。

4.4 值域要求

为方便数据上传单位数据录入、数据查询和数据库建立，字段描述的信息应首先考虑根据信息进行分类编码。分类编码须符合分类唯一性原则。

对于无法进行分类编码的信息字段，需说明该字段值域范围（如文本型的可定义为自由文本，日期型的需要定义时间范围）。

对于需要数据保密性的字段，对该字段数据进行定性描述，避免定量描述。

www.egene.org.cn
www.egene.org.cn

5 格式

5.1 标题

个性描述规范的主题。

5.2 前言

简单介绍规范的目的。

5.3 正文

5.3.1 范围

规定个性描述规范所包含的研究领域范围和适用范围。

5.3.2 规范性引用文件

列举规范引用文件。

5.3.3 定义

列举规范中的相关重要定义。

5.3.4 原则

规范编写的原则。

5.3.5 方法

编写的方法。

5. 3. 6 个性规范描述表

将个性描述规范的字段用进行分类，并以表格的形式展示。

www.egene.org.cn
www.egene.org.cn

人类遗传资源特性信息描述规范及编码标准字段（54 个字段）
… …
个性字段（X 个字段）
… …

5.3.7 个性描述规范简表

对个性描述规范的字段进行归类、排序，并给予简单的说明。其中人类遗传资源特性信息描述规范及编码标准字段的说明引用“人类遗传资源特性信息描述规范及编码标准字段”中相关的说明。

人类遗传资源特性信息描述规范及编码标准字段（54 个字段）
… …
个性字段（x 个字段）
… …

5.3.8 个性描述规范字段说明

描述规范字段说明分为两部分：

第一部分：引用“人类遗传资源特性信息描述规范及编码标准”中的字段说明；

第二部分：详细阐述个性字段所表示的内容，对值域可分类的字段给定其分类代码，分类代码表置于附录中（字段赋值代码表）。

5.3.9 附录 1：字段赋值代码表

分为两个部分：

第一部分：引用人类遗传资源特性信息描述规范及编码标准字段赋值代码表；

第二部分：列出个性描述规范中相关字段的赋值代码表。

5.3.10 附录 2：描述规范录入示例

按照前述的个性描述规范，具体给出一条数据录入记录，用以示例。

www.egene.org.cn

6 规范性附录：人类遗传资源特性信息描述规范及编码标准简表

护照信息			
平台资源号（1）		资源编号（2）	
内部编号（3）			
实物信息			
资源归类（4）		资源分类（5）	
样本类型（6）		样本定量（7）	
器官来源（8）		采集日期（9）	
保存条件（10）		保存期限（11）	
实物状态（12）		生物安全（13）	
资源用途（14）			
基本信息			
性别（15）		出生年月（16）	
籍贯（17）		民族（18）	
居住地（19）		职业（20）	
婚姻状况（21）		血型（22）	
生命周期（23）		文化程度（24）	
健康状况（25）			
特征信息			
鉴定资料（26）		干预资料（27）	
流行病学资料（28）		随访资料（29）	
家族资料（30）		疾病别名（31）	
其它资料（32）			
采集信息			
采集机构（33）		知情同意（34）	
保存单位（35）		采集设计（36）	
项目经费来源（37）		成果（38）	
备注（39）			
关联信息			
家系标记（40）		家系患者（41）	
组别标记（42）		病例对照（43）	
对象标记（44）		样本说明（45）	
图像信息			
图像（46）			
共享信息			
共享方式（47）		获取途径（48）	
联系单位（49）		邮政编码（50）	
联系电话（51）		联系人（52）	
E-mail（53）		共享利用信息（54）	

www.egene.org.cn