

通用公告

简介

中药材参考 DNA 序列库所收录的参考 DNA 序列，均源自政府中药检测中心(“检测中心”)研究项目中，分析已知分类位置的标本后所获得的数据。为了方便发布最新研究成果，检测中心会适时拟备资料表，向业界及公众发放所选用的 DNA 条形码之参考 DNA 序列及资料。

资料表内容编排及说明

每份资料表阐述一种中药材物种的资料及参考 DNA 序列，内有封面页、参考 DNA 序列和 FASTA 格式的参考 DNA 序列。

- (a) 封面页——概述所有用作产生数据的标本资料及参考 DNA 序列的特性，并会按需要提供产生个别生物 DNA 条形码的附加实验条件资料。
- (b) 参考 DNA 序列——提供每个标本的采集资料及参考 DNA 序列，编排顺序为 DNA 资料、参考资料和图像。
 - DNA 资料——按情况以 FASTA 格式列出附有专用序列识别号的预设 DNA 条形码。在大部分情况下，参考 DNA 序列都是利用检测中心内部验证的方法产生，整个相关工作流程已对外发布，可供自由参阅。所选用的条形码适用于鉴别物种，已获多国药典采用，或已得到科学界广泛认可。资料表载列每个标本多达三组 DNA 条形码，以增加鉴别物种的准确度及灵活度。
 - 参考资料——提供每个标本的详细资料，包括参考物料、学名、实验室编号、样品标记、产地、真伪、对应的凭证标本或中药材标本(如适用)及备注。
 - ◆ 所有参考 DNA 序列均来自三种来源已知的参考物料：生物凭证标本、经鉴定的生物部分或经鉴定的中药材。这些参考物料取自以下计划 / 组织所收集的留存样本：香港中药材标准(“港标”)计划、政府中药检测中心中药标本馆，以及公认的博物馆、国家机构、大学或研究机构(如适用)。
 - ◆ 每个标本和每种药材均经过专家鉴定，并由检测中心详细记录和永久保存。资料表所载标本的真确性、经验证的检测方法和 DNA 条形码的资料，为鉴别物种的遗传可追溯性奠定基础，保障中药材临床应用的安全和效用。

- ◆ 每份参考物料均经过两次独立的检测，确保所产生的参考 DNA 序列的质量。
 - ◆ 在可行的情况下，中药材参考 DNA 序列库会为采集自同一地域的相关凭证标本及其药材，于“相对应之凭证标本 / 相对应之中药材标本”一栏中标明其关系。
 - ◆ 所得数据如有特殊发现，则会加以注明。
 - 图像——显示标本的图片。
- (c) FASTA 格式的参考 DNA 序列——按 DNA 条形码划分，分别以 FASTA 格式列出所有标本的参考 DNA 序列，以便快速处理参考 DNA 序列，利便进行后续分析。

术语

- (a) 预设 DNA 条形码——指所选取用于特定生物界别的 DNA 条形码：
- 植物物种的预设 DNA 条形码包括核糖体 DNA 第二内部转录间隔区 (ITS2)、叶绿体 *psbA-trnH* 基因间区 (*psbA-trnH*) 及叶绿体二磷酸核酮糖羧化酶大链 (*rbcL*)。
 - 动物物种的预设 DNA 条形码包括线粒体基因组细胞色素 C 氧化酶亚基 I (COI)、线粒体基因组细胞色素 b (CYTB) 及线粒体基因组 16S 核糖体核糖核酸 (16SrRNA)。
- (b) FASTA 格式——指核苷酸或氨基酸序列的格式，是生物资讯学序列分析常用的数据格式。
- (c) 实验室编号——指标本的专用样品识别编码。
- (d) 核苷酸代码——指理论化学和应用化学国际联合会 (IUPAC) 使用的编码，表列如下：

IUPAC 核苷酸编码	译码
A	腺嘌呤
C	胞嘧啶
G	鸟嘌呤
T	胸腺嘧啶
B	C、G 或 T
D	A、G 或 T
H	A、C 或 T

IUPAC 核苷酸编码	译码
R	A 或 G (嘌呤)
Y	C 或 T (嘧啶)
K	G 或 T
M	A 或 C
S	G 或 C
W	A 或 T
N	A、C、G 或 T
V	A、C 或 G

- (e) 样品标记——指待测样品的专用识别标记，其最后的字母表示样品重复对照的编号，例如 RD100-A 和 RD201-1-A 分别是专用样品 RD100 和 RD201-1 的样品重复对照 A。
- (f) 序列识别号——指 FASTA 档案以“大于”符号（即‘>’）开首的参考序列名称。序列识别号由样品标记及预设 DNA 条形码的名称组成，例如序列识别号为“>RD100-A_ITS2”代表 ITS2 参考 DNA 序列来自样品标记为“RD100-A”的测试样品。

免责声明

- (a) 资料表所载的全部资料均视作政府财产，可用于个人、研究或教育用途，但受版权保护而不得用于转售或商业用途。
- (b) 文件内容如有变更，恕不另行通知。如因使用、误用或依据资料表所载的任何资料而导致任何损失或损害，政府中药检测中心概不负责。