

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2021/196356 A1

(43) 国际公布日
2021 年 10 月 7 日 (07.10.2021)

(51) 国际专利分类号:
G16B 30/10 (2019.01)

(21) 国际申请号: PCT/CN2020/090175

(22) 国际申请日: 2020 年 5 月 14 日 (14.05.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

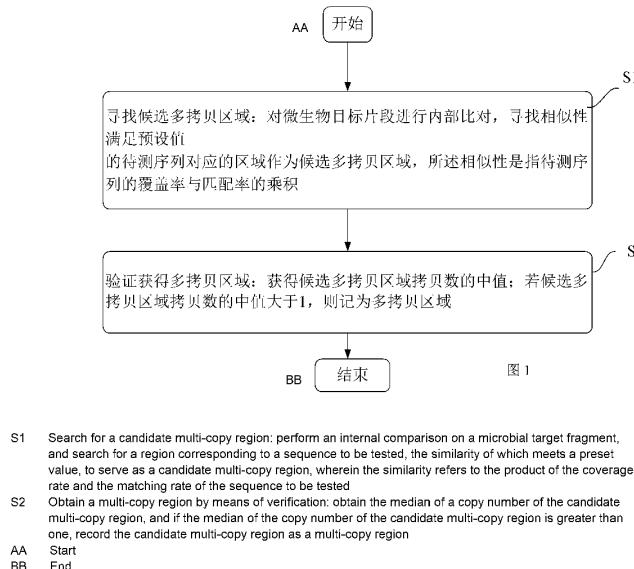
(30) 优先权:
202010254690.9 2020年4月2日 (02.04.2020) CN

(71) 申请人: 上海之江生物科技股份有限公司(SHANGHAI ZJ BIO-TECH CO., LTD.) [CN/CN];
中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。

(72) 发明人: 嵇匆(JI, Cong); 中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。邵俊斌(SHAO, Junbin); 中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。刘燕(LIU, Yan); 中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。齐霞(QI, Xia); 中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。金宇丹(JIN, Yudan); 中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。李启腾(LI, Qiteng); 中国上海市浦东新区张江高科技园区瑞庆路 528 号 20 幢乙号 1 层、21 幢甲号 1 层, Shanghai 201114 (CN)。

(54) Title: METHOD AND APPARATUS FOR IDENTIFYING MULTI-COPY REGION IN MICROBIAL TARGET FRAGMENT, AND USE

(54) 发明名称: 微生物目标片段中多拷贝区域的识别方法、装置及应用



(57) Abstract: A method for identifying a multi-copy region in a microbial target fragment. The method comprises at least the following steps: searching for a candidate multi-copy region: performing an internal comparison on a microbial target fragment, and searching for a region corresponding to a sequence to be tested, the similarity of which meets a preset value, to serve as a candidate multi-copy region, wherein the similarity refers to the product of the coverage rate and the matching rate of the sequence to be tested (S1); and obtaining a multi-copy region by means of verification: obtaining the median of a copy number of the candidate multi-copy region,

[见续页]

(74) 代理人：上海光华专利事务所（普通合伙）(J.Z.M.C. PATENT AND TRADEMARK LAW OFFICE (GENERAL PARTNERSHIP)); 中国上海市杨浦区国定路 335 号 5022 室余明伟, Shanghai 200433 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护)：AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

and if the median of the copy number of the candidate multi-copy region is greater than one, recording the candidate multi-copy region as a multi-copy region (S2). The method for identifying a multi-copy region in a microbial target fragment is high in accuracy and sensitivity; an undiscovered multi-copy region can be identified; repetitive sequences can be found in an incompletely assembled motif; and compared with a method using 16S rRNA, the method is more comprehensive.

(57) 摘要：一种微生物目标片段中多拷贝区域的识别方法，至少包括以下步骤：寻找候选多拷贝区域：对微生物目标片段进行内部比对，寻找相似性满足预设值的待测序列对应的区域作为候选多拷贝区域，所述相似性是指待测序列的覆盖率与匹配率的乘积（S1）；验证获得多拷贝区域：获得候选多拷贝区域拷贝数的中值；若候选多拷贝区域拷贝数的中值大于1，则记为多拷贝区域（S2）。该识别方法的微生物目标片段中多拷贝区域的识别方法准确性高，灵敏度高，识别出未发现的多拷贝区域；可在组装不完整的基序中寻找重复序列；与16srRNA相比更加全面。