**审稿意见与作者修改说明（稿号：2020-0154）**

——————————————**初审专家意见与作者修改说明**—————————————

专家1修改意见及回复：

1.“咽部？口咽部？哪个更准确一点？请推敲。”

回复：感谢专家的提出的宝贵意见，同意专家意见，修改稿行号27中“口咽部”已更正为“鼻咽部”；

2.“请对Nem缩写进行注释。”

回复：感谢专家的提出的宝贵意见，同意专家意见，修改稿行号125-126中缩写已更正为：A群流脑菌株(NmA)、X群流脑菌株(NmX)。

专家2修改意见及回复：

“1：作者应将本研究所用菌株/基因组的详细信息用表格的形式总结出来，以方便读者更好的理解试验结果；”

回复：感谢专家的提出的宝贵意见，同意专家意见，已将本研究所用菌株/基因组的详细信息总结在附表1中，并在修改稿行号46-47中更正。

“2：本研究所用菌株/基因组的血清群型别是如何确定的？血清凝集或基于序列分析？”

回复：感谢专家的提出的宝贵意见，同意专家意见，本研究所用菌株/基因组的血清群型基于血清凝集的方法，已在修改稿行号48-50中更正。

“3：如何评价本研究所发现的荚膜基因簇呈现的多态性的生物学意义？本研究所确定的荚膜基因簇呈现的多态性是否导致相应菌株表型的多态性？作者是否开展过相关的实验研究？结果如何？如果没有开展相关的研究，作者应该在讨论部分里引用相关文献展开说明。”

回复：感谢专家的提出的宝贵意见，同意专家意见。已在修改稿行号132-135的讨论中引用相关文献展开说明，课题组也会在后期大数据研究中关注和分析荚膜基因簇多态性与菌株表型的多态性之间的关系。

“4：作者使用Pairwise distance值（P-distance）来表示基因间的序列多态性，建议作者增加相关基因间的blast比对结果（包括identity和coverage值），方便读者更好的理解；基因序列的差异作者仅做了简单的描述，应做更详细的归纳与总结，如序列存在变异的基因在CPS合成中的作用是什么？变异可能会造成表型的何种改变？序列变异有无规律，有无ST型/血清群型别相关性？等等”

回复：感谢专家的提出的宝贵意见，同意专家意见。现已将各基因长度和用来描述基因间多态性的P-distance值补充到附表2，并在修改稿行号61、63中标注。

附表2中显示相同血清群基因间长度大多数相同，个别基因存在5bp或者6bp，因此并未进一步显示coverage值。由于C血清群比B血清群的cps多一个cssE基因，并且csb和csc基因是B群和C群的特异性基因，为了表示基因间的序列多态性，文章选用了准确且操作便捷的MEGA软件中的Pairwise distance值（P-distance）来表示多态性。为了便于读者更好理解各区基因在CPS合成中的作用，文章补充了在附图1 NmB和NmC的cps基因分布和功能图，并在修改稿行号123-125中补充了变异的基因在CPS合成中的作用，以及哪些变异基因是血清群的特异性基因。目前尚未见序列变异与ST型的关系，亦未见相关报道，课题组将在后续大数据研究中进一步分析和归纳两者关联性。

“5：作者发现了荚膜基因簇序列中存在较多的序列重组情况，这些重组序列两端是否存在反向重复序列？是否有重组酶参与？重组的机制如何？作者应在结果和讨论中加以说明。”

回复：感谢专家的提出的宝贵意见，同意专家意见。重组机制与反向重复序列的关系以及重组机制的研究进展，已修改稿行号128-131补充说明，并且补充了41条Nm cps A\*CE、A\*、C和E区Neighbor joining系统发育树（附图2），以方便读者更好理解重组事件的鉴定过程。

“6：cps各区的D-A-C-E-D’-B的排布方式发生改变的生物学意义是什么？作者应结合文献在讨论中加以说明。”

回复：感谢专家的提出的宝贵意见，同意专家意见。已修改稿行号147-150补充说明。

“7：作者在本文中同时使用nt和kb来表示基因长度与位置信息，作者应使用统一的标准。建议使用“bp”和“kb”来表示。”

回复：感谢专家的提出的宝贵意见，同意专家意见，已将修改稿及图表中基因长度与位置信息统一更正为“bp”和“kb”

“8：本研究涉及的基因序列是否有公开数据库的accession No？如有需要注明。”

回复：感谢专家的提出的宝贵意见，同意专家意见，本研究涉及的24条基因序列有公开数据库的accession No.，已在附表1中标注，获得的GenBank ID为MN683839-MN683862，已修改稿行号46-47中更正。

——————————————**复审专家意见与作者修改说明**—————————————

专家意见：

1. “能否统一克隆群的名称？还有后面的CC4821”

回复：感谢编辑的宝贵建议，同意编辑建议。文中标题和文中第一次出现使用标准全称“ST-4821克隆群”，其余统一改为“CC4821”

2. “补充具体试剂和公司全称”

回复：感谢编辑的宝贵建议，同意编辑建议。已将句子更正为“按照Remel 公司的Nm 血清凝集试剂盒说明书进行菌株血清群复核，按照Promega 公司的DNA 提取试剂盒说明书进行核酸提取”

3. “第二个菌株编号是不是写错了？这个地方是不是可以直接写以AM421808.1和CP000381.1为参考”

回复：感谢编辑的宝贵建议，同意编辑建议。已将文中改为“AM421808.1和CP000381.1”

4. “请核实无误”

回复：感谢编辑的宝贵建议，同意编辑建议。已核实，并将文中句子改为“命名为A\*区”以方便读者理解

5. “大写或小写，全文保持统一”

回复：感谢编辑的宝贵建议，同意编辑建议。已将文中统一更正为“p-distance”

6. 图1中“（1）图中各项名称改为中文，如strain、region、year等；（2）是否需要年的信息？（3）图中左侧用GenBank号还是菌株号？菌株号和血清群能否分开写？请核实哪个更合适；（4）删除“Scale （kb）”，在刻度15后加kb；（5）删除图的外边框；（6）删除图例的外边框和其中的：和；”

回复：感谢编辑的宝贵建议，同意编辑建议。（1）已完成；（2）年份信息已删除；（3）图中与正文统一使用菌株编号，并已修改图注。图中血清群已分开写；（4）已删除“Scale （kb）”，并在刻度15后加kb；（5）已删除图的外边框；（6）已删除图例的外边框和其中的：和；。

7. 图2及注释已参照图1修改

8. “图3太小，完全看不清楚”

回复：感谢编辑的宝贵建议，同意编辑建议。已参照图1编辑建议修改图3，同时，已尽量放大图3并且将B区字体改为红色，方便读者理解图3表达的意思。此外，由于基因过多，基因名称可能看不清楚，所以，图3将“B区及附近基因”基因用不同的颜色表示，方便读者理解。

9. “参考文献20先于19”

回复：感谢编辑的宝贵建议，已经将遗漏的19补充上，文中更正为“[9, 13, 19-20]”

10. 两处不通顺句子和建议修改讨论部分的句子

回复：感谢编辑的宝贵建议，已更正两处不通顺句子，改为“断点位于*cps*的2026bp和6270bp位点”，“ACE区多态性主要体现在*ctrD*、*ctrC*和*tex*这三个基因”。讨论部分已经通读修改，并在文中更正。

专家2:

1.“图1：（1）基因应该采用标准的箭头标示，不应该是方框加箭头；（2）将ctrG基因左端对齐；（3）刻度尺放图的最上方；（4）把菌株编号和血清型放在“FAM18 C”的上方；（5）删除“基因”、“分区””

回复：感谢编辑的宝贵建议，同意编辑建议。图1中已按照上述建议完成修改。

2. “保守、差异大、差异小的数值具体什么？请补充”

回复：感谢编辑的宝贵建议，同意编辑建议，正文中已补充相关数值。

3. “请核实参与下文分析的菌株数量”

回复：感谢编辑的宝贵建议，同意编辑建议。已核实分析参与菌株数量，文中共分析实验菌株24株（16株NmB、8株NmC）和17株的全基因组序列，共计41株。附表1和附图2中显示41株菌株的背景信息和进化分析。为了使文章简洁，方便读者理解，文章结果只重点展示具有新发现的菌株，并使用参考菌株和2株未见新发现的实验菌株作为比对，以方便读者理解。

4. 图2“将重组模块与所分析菌株的示意基因组叠加，尽量减小序列间的间隔；刻度尺放图的最上方；把菌株编号和血清型放在“FAM18 C”的上方；删除“基因”、“分区””

回复：感谢编辑的宝贵建议，同意编辑建议，已在图中修改

5. “分别介绍D和D’区的长度”

回复：感谢编辑的宝贵建议，同意编辑建议，已在正文中补充。

6. 表格1“根据正文做相应修改，只体现核苷酸和氨基酸的相似度，不写差异位点数”“名称与正文保持一致”“这一列数字不对，请核实。galE2和galE的长度经常不一致，确定这些菌株左右重复区序列长度一致？”

回复：感谢编辑的宝贵建议，同意编辑建议，已修改表格1和正文中内容。

Self-blast 是把一个菌株的cps 序列和它自己做比对，找出重复的区域，所以会得出2个重复区，一个位于D区，一个位于D'区，然后进下一步分析这2个重复区的差异。所以，两个重复区长度是一致的。

表1中先分析重复区核苷酸差异性，再进一步分析重复区rfbC/rfbC2、rfbA/rfbA2、rfbB/rfbB2和galE/galE2氨基酸差异性，更好地凸显“D和D’区有明显差异序列”这个结果，方便读者理解。

7. “分析了多少株菌的序列？”

回复：感谢编辑的宝贵建议，同意编辑建议。正文已更正为：“通过分析41条cps序列，本研究发现8条cps的B区位置发生了变化（菌株编号为131457、210713、310827、320701、321114、360624、431002和421615）”，并添加了8条cps菌株编号方便读者理解和参考。

8.图3“做成类似图1的示意图；如果要展示两侧外围基因，请在所有菌株中统一展示：

回复：感谢编辑的宝贵建议，同意编辑建议。已按照建议修改图3.