

mNeonGreen 荧光蛋白的定向进化

凌孝语

(麦吉尔大学 加拿大 蒙特利尔 H3A0G4)

【摘要】mNeonGreen 是一种黄绿色荧光蛋白，是现有的最亮的单体荧光蛋白。它源自四聚体 LanYFP，这是已知最亮的天然荧光蛋白。LanYFP 起源于头索动物 *Branchiostoma lanceolatum*，通过诱变定向进化成为 mNeonGreen。由于其令人印象深刻的荧光特性，mNeonGreen 在充当体内成像和 FRET 的融合标签等任务中表现出色。然而，它的特性可以进一步扩展，以适应更广泛的生物研究领域。该项目旨在使用易错 PCR (EP-PCR) 在野生型 mNeonGreen 基因中产生随机点突变，以产生具有改进特性的新版本蛋白质。由此产生的突变体 mNeonGreen 基因包含一个过渡突变，导致荧光蛋白外部桶状结构上的极性氨基酸被疏水性氨基酸取代，同时使发色团区域保持不变。这导致蛋白质稳定性较差，因为表达突变体 mNeonGreen 的大肠杆菌的光漂白半衰期减少。表达突变体 mNeonGreen 的大肠杆菌的成像样本显示，由于荧光蛋白浓度较高，荧光大大增加。与野生型 mNeonGreen 相比，突变型 mNeonGreen 的荧光分光光度分析显示荧光强度、pH 敏感性 (pKa) 或激发 / 发射光谱没有差异。因此，研究人员假设涉及 mNeonGreen 发色团区域周围区域的突变更有可能产生 mNeonGreen 变体。将来，也许定点诱变可能有助于特异性靶向 mNeonGreen 的发色团区域，以产生新的变体。

【关键词】 荧光蛋白；大肠杆菌；靶向突变；荧光强度；基因突变

The Directed Evolution of mNeonGreen Fluorescent Protein

Xiaoyu Ling

(McGill University, Montreal, Canada, H3A0G4)

[Abstract] mNeonGreen is a yellow-green fluorescent protein, the brightest monomeric fluorescent protein available. It was derived from the tetrameric LanYFP, the brightest natural known fluorescent protein. LanYFP originated from the cephalochordate *Branchiostoma lanceolatum*, and it underwent directed evolution through mutagenesis to become mNeonGreen. Due to its impressive fluorescent properties, mNeonGreen performs exceptionally well in tasks such as acting as a fusion tag for in-vivo imaging and FRET. However, its properties can be further expanded to suit an even wider range of biological research areas. This project aims to use error prone PCR (EP-PCR) to generate random point mutations in wild-type mNeonGreen gene to produce new versions of the protein with improved properties. The resulting mutant mNeonGreen gene contained one transition mutation, causing a polar amino acid on the outside barrel structure of the fluorescent protein to be replaced by a hydrophobic amino acid while leaving the chromophore region untouched. This resulted in a less stable protein as the photobleaching half-time of *E. coli* expressing mutant mNeonGreen was reduced. Imaging samples of *E. coli* expressing mutant mNeonGreen showed that fluorescence greatly increased due to higher concentrations of fluorescent proteins. Spectrofluorometry analysis of mutant mNeonGreen showed no differences in the fluorescence intensity, pH sensitivity (pKa), or excitation/emission spectra when compared with wild-type mNeonGreen. Therefore researchers hypothesize that mutations involving areas around the chromophore region of mNeonGreen is more likely to produce mNeonGreen variants. In the future, perhaps site directed mutagenesis may help to specifically target the chromophore region of mNeonGreen to generate new variants.

[Key words] Fluorescent protein; *Escherichia coli*; Targeted mutation; Fluorescence intensity; Gene mutation

1 目的

为了产生突变型 mNeonGreen 基因，进行了 EP-PCR 以便在野生型 mNeonGreen 基因中随机突变。为了产生突变的 mNeonGreen 蛋白，将突变基因插入载体，克隆到大肠杆菌细胞中，并诱导蛋白表达。为了表征突变体 mNeonGreen 蛋白的特性，使用生物信息学分析基因和蛋白序列，进行蛋白纯化以生成纯化的突变体

蛋白，进行 ChemiDoc 成像以粗略分析荧光，并进行 SDS-PAGE/Western blot 来分析纯化效率以及突变体 mNeonGreen 在大肠杆菌细胞中的表达水平。进行荧光光谱分析以确定突变体荧光强度、激发 / 发射光谱和 pH 敏感性。对含有突变 mNeonGreen 蛋白的活大肠杆菌细胞进行基于荧光显微镜的分析，以确定突变体光漂白半衰期。

2 材料与方法

2.1 克隆验证和测序

5 管 5mL 液体 Luria Broth (LB) 培养物 (标记为 MP), 无细胞 (NO)、带有质粒 pUC18-His 且未插入基因的大肠杆菌 (H)、插入野生型 mNeonGreen (C) 或插入突变型 mNeonGreen (M), 培养过夜 (振荡培养箱, 37℃)。4 管 1mL LB 培养物 (标记为 PE+H、C 和 M) 也培养过夜 (转鼓, 37℃)。对 MP 管进行基于硅胶的质粒纯化, 纯化后, 将 3 μ L ddH₂O 和 2 μ L 纯化的 DNA 样品与 15 μ L KpnI 或 KpnI+XmnI 限制性酶混合物 (对于所有样品) 混合并孵育 (37℃水浴, 30 分钟)。然后将 2 μ L 染料与 20 μ L 消化样品混合, 每种样品总共 20 μ L 与 GeneRuler™ 一起上样到琼脂糖凝胶上进行凝胶电泳 (120V, 约 45 分钟)。使用 ChemiDoc 对凝胶进行成像。最后, 将 15 μ L 带有突变 mNeonGreen 基因的纯化质粒送至 Genome Québec 进行 Sanger 测序。

2.2 序列和结构分析

使用了生物信息学工具。BLASTn 对野生型 mNeonGreen 基因与突变型基因的核苷酸序列进行比对和比较。ExPASy 将 mNeonGreen 基因的核苷酸序列翻译成蛋白质的氨基酸序列。BLASTp 比对并比较了野生型 mNeonGreen 蛋白与突变蛋白的氨基酸序列。蛋白质数据库分析了 mNeonGreen 蛋白的二级结构, Swiss Model 进行了模板比对模型, PyMol 构建了突变 mNeonGreen 蛋白的 3D 模型进行分析。

2.3 蛋白质纯化

克隆验证和测序的 PE 管 (H、C、M) 加入 9mL 液体 LB+ 氨苄青霉素, 培养 (37℃, 振荡培养箱, 4 小时) 至指数生长期。在本生灯营造的无菌环境下, 每管中添加 5 μ L IPTG (1M) 以诱导蛋白表达。将这些管再次生长 (37℃, 振荡培养箱, 再培养 4 小时)。然后将每管中的 2mL 离心, 弃去上清液并保留沉淀 (重复 4 次, 因此大肠杆菌沉淀来自总共 8mL)。对每个样品的总蛋白提取物进行 1000 倍稀释 (前 10 μ L 样品用 990 μ L ddH₂O 进行 100 倍稀释, 然后用 720 μ L ddH₂O 稀释 80 μ L 100 倍)。还创建了每个样品的纯化蛋白的 10 倍稀释液 (80 μ L 纯化蛋白与 720 μ L ddH₂O)。将稀释液涡旋混合。Bradford 测定用于测量蛋白质浓度 (Kielkopf 等人, 2020)。将 200 μ L Bradford 试剂 (考马斯亮蓝 G-250) 与 800 μ L 稀释的蛋白质样品 (总量 1000 倍, 纯化 10 倍) 混合并孵育 (10 分钟)。使用分光光度计测量样品在 595nm 处的光密度以计算蛋白质浓度。

通过 SDS-PAGE 和蛋白质印迹进行蛋白质分析。

使用 SDS-PAGE 分析蛋白质纯化效率 (Brunelle

和 Green, 2014)。

3 结果

3.1 蛋白质序列和结构分析

桑格测序结果提供了突变体 mNeonGreen 基因的核苷酸序列。ExPASy 翻译提供了突变体 mNeonGreen 蛋白的氨基酸序列。BLASTn 对比野生型和突变型 mNeonGreen 基因序列后, 结果表明 e 值为 0.0, 比对中不存在缺口。在查询序列的第 151 位, 735 个核苷酸中有 1 个差异, 并且突变基因中的胸腺嘧啶 (T) 突变为胞嘧啶 (C)。这是一种过渡突变 (嘧啶到嘧啶)。ExPASy 翻译后, 选择最长的阅读框进行分析。BLASTp 对比野生型和突变型 mNeonGreen 蛋白序列后, 结果表明 e 值为 0.0, 比对中不存在缺口。查询序列第 51 位的 244 个氨基酸中有 1 个存在差异, 并且突变蛋白序列中的丝氨酸 (S) 变为脯氨酸 (P)。丝氨酸是极性氨基酸, 脯氨酸是非极性氨基酸。模型模板比对表明该氨基酸突变是 β 折叠二级结构的一部分 (图 1A)。突变体 mNeonGreen 的 3D 模型可以可视化其结构变化的位置, 以红色 (脯氨酸) 突出显示 (图 1B)。

3.2 荧光强度的表征

在 495nm $\pm$ 10nm 激发和 520nm $\pm$ 10nm 发射设置下进行荧光分光光度计分析后, 野生型 (C)mNeonGreen 的荧光强度 (归一化, 减去背景强度) 为 471.087au。突变体 (M) 的荧光强度为 451.431au, 与野生型相比, 荧光强度为 95.83% (图 2A)。对于激发/发射光谱, 突变体的曲线与野生型 (两者) 非常相似 (图 2C、4D)。对照的峰值激发/发射波长为 504nm 和 517nm, 这与报道的 mNeonGreen 在 506nm 激发和 517nm 发射处的峰值非常相似 (Shaner 等人, 2013)。突变体的峰值也相似, 在 505nm 激发和 517nm 发射 (图 4B)。将蛋白质暴露于不同 pH 溶液后, 使用分光荧光计测量荧光强度, 以计算其对 pH 的敏感性。绘图后, 可以使用线性方程生成的方程来估计每种蛋白质的 pKa。对于野生型, $y=42.021x-176.79$, 对于突变体, $y=48.755x-212.68$, 其中 y 是荧光强度 (%), x 是 pH (图 2E)。野生型和突变体之间的曲线基本相同, 因为荧光强度随着 pH 值达到 7.5 而变高。不同之处在于突变体在 pH6.5 时具有最高荧光强度, 而不是野生型在 7.5 时 (图 2E)。pKa 估计为荧光达到最大值一半 (50%) 时的值。mNeonGreen 的报告 pKa 为 5.7 (Shaner 等, 2013), 野生型的计算 pKa 为 5.40, 突变体的计算 pKa 为 5.39。

4 讨论

BLASTn 后, e 值为 0, 小于 1×10^{-10} , 表明突变体的核苷酸序列与野生型匹配良好 (e 值显著, 比对和相

似性不是偶然的)。序列之间的起点和终点也相似,表明它在同一阅读框中。没有间隙是序列对齐良好的另一个迹象。突变(T到C)是一种过渡突变(嘧啶到嘧啶),通常比颠换突变危害更小,因为过渡突变更有可能在蛋白质水平产生沉默突变。经过 ExPASy 翻译和 BLASTp 分析后, e 值也为 0,表明它是显著的(小于 1×10^{-4})并且没有间隙。丝氨酸是极性的,脯氨酸是疏水性的。极性氨基酸往往位于蛋白质的表面,以稳定其与亲水环境(水、氢键)的相互作用,而疏水性残基往往位于核心。事实上,这种突变将桶上的残基从极性改变为疏水性,这一事实很可能会影响细胞内蛋白质的稳定性,因为与环境的相互作用将被破坏。发色团的结构没有改变,这意味着蛋白质的荧光特性很可能不会受到影响。

对于 SDS-PAGE, 每个样品上样等量的蛋白质,以确定哪个样品实际上表达更高或更低量的 mNeonGreen (使用总提取泳道)。纯化的 His 样品不含条带,因为它是阴性对照。它不应该含有 mNeonGreen 蛋白质,因此纯化后它根本不含蛋白质。尽管 His 样品含有组氨酸聚合物,但它们非常小,很可能会丢失在凝胶中,这意味着它们无法检测到。根据梯子,总野生型(C)、总突变体、纯化野生型(C)和纯化突变体中的突出条带约为 26kDa。实际 mNeonGreen 的分子量约为 26.6kDa (Shaner 等人, 2013), 显示与这些条带的大小相关性。纯化的 His 泳道没有条带,而野生型/突变体泳道只有 1 个条带,这一事实强烈表明该条带对应于 mNeonGreen。这意味着 mNeonGreen 的选择性纯化步骤是成功的。野生型和突变体的纯度%相似,但突变体的高产量是由于高蛋白质浓度(之前讨论过)。Western 转移是成功的,因为转移后,凝胶现在没有条带,而膜有条带图案,这

意味着蛋白质转移到膜上。Western blot 检测条带也对应于 mNeonGreen, 因为一抗对 mNeonGreen 蛋白上的 His 具有选择性。当与 Opti-4CN 底物一起提供时, HRP 会产生沉淀。突变体的高信号密度还归因于高蛋白产量(较高的 mNeonGreen 表达),这是由于纯化后的高蛋白浓度(之前讨论过)。

以上所有结论纯属猜测,还需要进一步调查证实。mNeonGreen 上的这种特殊突变似乎极大地增强了细胞中蛋白质的表达,但同时似乎也降低了蛋白质的稳定性,使其更容易发生光漂白。为了确认其影响,需要对这种特定突变进行进一步研究。这种突变也没有影响 mNeonGreen 的荧光特性。先前已确定荧光蛋白的发色团区域负责荧光特性,这意味着针对发色团区域的定点诱变可能更有效地产生具有不同荧光特性的突变体(Clavel 等人, 2016)。mNeonGreen 已经是一种由 LanYFP 定向进化产生的蛋白质。该过程涉及发色团附近的两个显著突变,导致荧光强度损失 35%,紫外-可见吸收和荧光发射最大值均发生 7nm 蓝移(Clavel 等人, 2016)。基于此,可以合理地假设,专门针对 mNeonGreen 发色团区域周围区域的突变(使用位点特异性诱变而不是随机诱变)可能会增加产生具有不同荧光特性的新突变体的机会。

由于 mNeonGreen 已经是一种具有强荧光的蛋白质,也许本实验室所做的工作可以适应于增强其他与 mNeonGreen 相比荧光较弱的蛋白质的蛋白质特性。使用针对发色团区域的定点诱变,可以改善其他较弱荧光蛋白的荧光特性。

分析 mNeonGreen 蛋白质序列和结构

A. 野生型 mNeonGreen 和突变型 mNeonGreen 之

实验图片

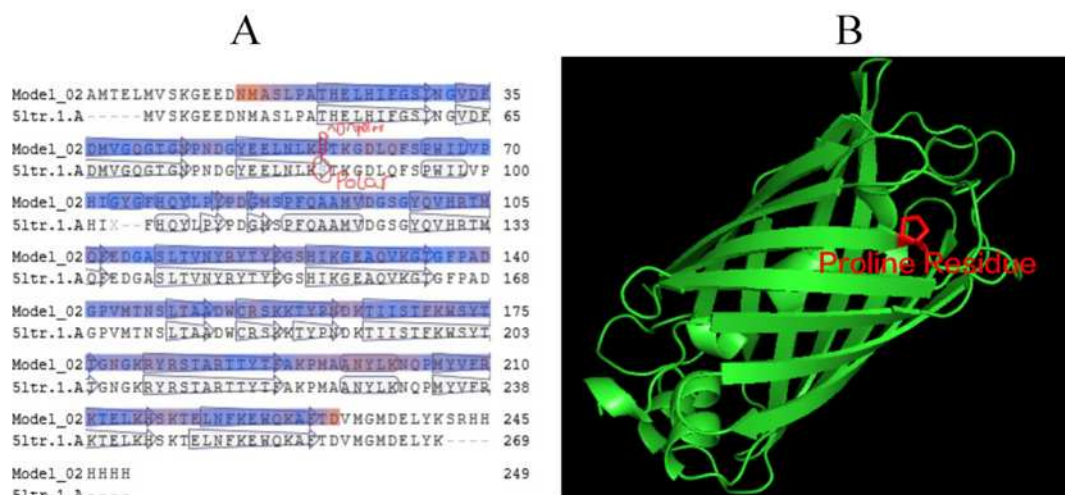


图 1. 使用生物信息学工具

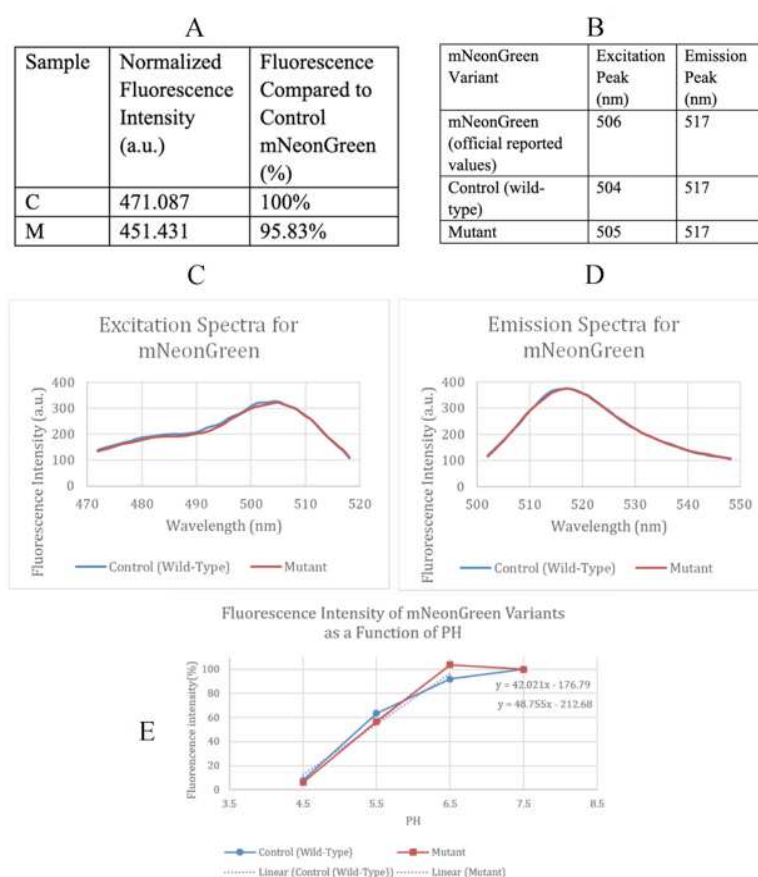


图 2. 荧光强度、激发 / 发射光谱的差异，野生型和突变型 mNeonGreen 之间的 pH 和 pH 敏感性

间的蛋白质氨基酸序列比对。显示了二级蛋白质结构，箭头是 β 折叠，椭圆形是 α 螺旋。该突变体在第 51 位包含疏水性脯氨酸 (P)，而不是亲水性丝氨酸 (S)。该突变是 β 片层结构的一部分。

B. 3D 模型可可视化突变 mNeonGreen 蛋白的结构，从而可视化突变残基的位置。脯氨酸残基清晰可见，因为与蛋白质其余部分的绿色相比，它以红色突出显示。可以清楚地看到氨基酸脯氨酸的环结构。

参考文献：

- [1] Brunelle, J.L., and Green, R. (2014). One-dimensional SDS-Polyacrylamide Gel Electrophoresis (1D SDS-PAGE). *Methods in Enzymology*, 151.
- [2] Clavel, D., Gotthard, G., Von Stetten, D., De Sanctis, D., Pasquier, H., Lambert, G.G., Shaner, N.C., and Royant, A. (2016). Structural analysis of the bright monomeric yellow-green fluorescent protein mNeonGreen obtained by directed evolution. *Acta Crystallographica Section D Structural Biology* 72, 1298–1307. 10.1107/s2059798316018623.
- [3] Hnasko, T.S., and Hnasko, R.M. (2015). The Western Blot. In *Methods in Molecular Biology*, (Springer

New York), pp. 87–96. 10.1007/978-1-4939-2742-5_9.

- [4] Hostettler, L., Grundy, L., Käser-Pébernard, S., Wicky, C., Schafer, W.R., and Glauser, D.A. (2017). The bright fluorescent protein mNeonGreen facilitates protein expression analysis in vivo. *G3: Genes, Genomes, Genetics* 7, 607–615.

- [5] Jacobson, G., and Kårnsås, P. (1990). Important parameters in semi-dry electrophoretic transfer. *Electrophoresis* 11, 46–52. 10.1002/elps.1150110111.

- [6] Kielkopf, C.L., Bauer, W., and Urbatsch, I.L. (2020). Bradford Assay for Determining Protein Concentration. *Cold Spring Harbor Protocols* 2020, pdb.prot102269. 10.1101/pdb.prot102269.

- [7] Remington, S.J. (2011). Green fluorescent protein: A perspective. *Protein Science* 20, 1509–1519. 10.1002/pro.684.

- [8] Shaner, N.C., Lambert, G.G., Chammass, A., Ni, Y., Cranfill, P.J., Baird, M.A., Sell, B.R., Allen, J.R., Day, R.N., Israelsson, M., et al. (2013). A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nature Methods* 10, 407–409. 10.1038/nmeth.2413.