Date: 10/31/2017
Document Version: 1.0

Notes On Plasmids In This Collection

The plasmids in this collection were used for the work presented in (1). Many of the specific details regarding these plasmids and the fluorescent proteins (FPs) they encode can be found in the main text and supplemental information of our publication. This document highlights some of the key features which we hope will be useful for anyone interested in using these plasmids.

If you use the FPs from this collection, besides citing our work (1), please also include a citation to the original researchers who created these FPs. For your convenience, Table 1 of this document includes all citations to those sources.

(1) VECTORS

The pEB1 vector is used for all FPs derived from *Aequorea victoria* GFP, i.e.the cyan, green (uv-excitable), green and yellow-green FPs. The pEB1 vector derives from the low-copy plasmid pUA66 (2). pUA66 contains an origin of replication which is a variant of the origin from pSC101. pEB1 was constructed by inserting a constitutive promoter proC (3) upstream the ribosomal binding site (RBS) of pUA66 (derived from the T7 phage Gene 10 RBS). For each plasmid, the coding sequence for the FP of interest was inserted in place of *gfpmut2* in pUA66.

The pEB2 vector is used for all orange-reds, reds and far-red FPs. In pEB2, the origin of replication of pEB1 was replaced with the original wild-type origin of replication from pSC101 (4, 5).

(2) CODING SEQUENCE OF FLUORESCENT PROTEINS

All FPs in the collection derive from (1) GFP from *Aequorea victoria* (avGFP/wtGFP), (2) red fluorescent protein from *Discosoma sp.* (DsRed), and (3) fluorescent proteins eqFP578 and (4) eqFP611 from *Entacmaea quadricolor* eqFP578. Table 1 summarizes this relationship.

As a guiding principle, we first defined a coding sequence for each ancestor FP. For every FP related to each ancestor, we retained the same coding sequence except in locations where mutations specific to the FP occur. For example, the coding sequence of mEYFP is identical to the coding sequence of avGFP except at S65G, V68L, S72A, T203Y, and A206K.

All mutations with respect to the ancestor sequences have been annotated in a genbank format file which is available for each plasmid on their individual Addgene entry. Below, we summarize several important features.

(i) Defining a Coding Sequence for Ancestral FPs

Below, we describe how the "reference" coding sequences were defined.

Ancestral FP	**Coding Sequence Source**	**Comments**
avGFP (wtGFP)	GenBank: M62653.1	Original coding sequence from *A. victoria*
DsRed	pZS2-123 (6)	Inferred from mCherry coding sequence in pZS2-123 using arbitrary codons where needed.
eqFP578	GenBank: EF606900.1	Inferred from TagRFP coding sequence using arbitrary codons where needed.
eqFP611	N.A.	mRuby3 coding sequences derive directly from Addgene plasmid #74252

(ii) Second Amino Acid Residue

Following the original wtGFP amino acid sequence, our constructs do not have a valine (codon GTG) usually added at the second position of the amino acid sequence. This valine was originally inserted to increase translation efficiency in eukaryotes (7). For a few select FPs, we generated variants that include the 2nd position valine. These variants can be found on plasmids that contain a "2ndVal" designation.

(iii) Codon Optimized Variants

Our collection also includes some codon optimized FPs (coFPs). FPs in our collection with a "co" prefix (e.g. coSCFP3A) were derived from previously published *E. coli* codon optimized cyan and yellow FP sequences (6).

(3) NOTES ON SPECIFIC FLUORESCENT PROTEINS

(i) moxFPs

The initial description of moxFPs included neutral mutations H231L and Q80R (2). We did not include these mutations in our constructs.

(ii) mCerulean and mCerulean ME

For mCerulean, we used the sequence defined in (3, 4) given that the definition in the original publication (5) is ambiguous: Cerulean is defined as (ECFP/S72A/Y145A/H148D) with the definition of ECFP given in reference (6). However, if we use an earlier definition of ECFP (7), mCerulean would contain an extra mutation, N164H, not included in the modern version of mCerulean (3, 4). mCerulean ME contains this N164H mutation.

(iii) mEmerald

The version of mEmerald in our collection follows the version described in (2). This version differs from the original mEmerald sequence by one mutation, F64L (8).

(iv) Venus NB, Venus JBC, Venus ME and Venus SX

During our work, we discovered that there were four different variants of "Venus", each with slightly different non-synonymous CDSs. We named these variants Venus NB(8), Venus JBC(9), Venus ME(6), and Venus SX(10).

We call Venus NB the original Venus FP described by Miyawaki *et al.*(9). When first published, the fast *in vitro* refolding/oxidation of Venus NB was highlighted. In a subsequent study, Miyawaki *et al.* reported the crystal structure of Venus(10). In that study, they added a new substitution, V68L; we call this variant Venus JBC. Four years later, Kremers *et al.*(3) took Venus JBC and reverted L68 back to V68 to get Venus NB again. Kremers *et al.* called this revertant SYFP2. mVenus ME is mVenusJBC with the extra mutation Q69M. Venus SX is Venus JBC with extensive substitutions in the C-terminus (9).

(v) mCherry-L and mCherry2-L

The original development of mCherry involved appending the N- and C- termini of avGFP to a core sequence derived from mRFP1. During this process, the initial methionine of mRFP1 was retained. Subsequently, it was shown that this downstream methionine can act as an alternative translation start site (17). We decided to eliminate translation from the second start site by mutating the methionine into a leucine; we indicated this change by appending an "L" to the name of the FPs: mCherry-L or mCherry2-L.

(vi) mRFP1 and mRFP1*

mRFP1 is the regular mRFP1 and mRFP1* is mRFP1 plus the N- and C-termini of avGFP.

(4) SUGGESTIONS FOR CLONING AND SEQUENCING.

To amplify the cassette comprising the constitutive promoter and FP, we recommend the following primers.

T1_from_mSC101_R02	AAAGGGAAAACTGTCCATATGCAC
pZSeq2_Kan	AGGGCTTCCCAACCTTACCAG

Expected size from pEB1 ~1.6 kb (1585 bps for most plasmids)

Expected size from pEB2 ~1.1 kb. (1130 bps for most plasmids)

You can use primer pZ_seq_frw to sequence the cassette obtained from the previous pair of primers.

pZ_seq_frw	GCGTATCACGAGGCCCTTTC

These primers are annotated in all the constructs.

Ancestor	avGFP (wtGFP)	DsRed	eqFP578	eqFP611
Related FPs	mCerulean (11-13) SCFP1 (11) SCFP3A (11) mCerulean ME (6) mTurquoise (14) mCerulean3 (12) mTurquoise2 (15) moxCerulean3 (16) Sapphire (17) T-Sapphire (18) mEGFP (7) mGFPmut2 (19) mGFPmut3 (19) mEmerald (20) sfGFP (21) moxGFP (16) mEYFP (20) mVenus NB (8) mVenus JBC (9),(11) mYPet (22) mVenus ME (6) Venus SX (10) Clover (23) moxVenus (16) mClover3 (24)	DsRed-Express (25) DsRed-Express2 (26) mRFP1 (27) mRFP1* (27) mScarlet (28) mScarlet-I (28) mCherry-L (29) mCherry2-L (30) E2-Crimson (31)	TurboRFP (32) TagRFP (32) TagRFP-T (33) mKate2 (34) mNeptune2 (35) mNeptune2.5 (35) Katushka (36) Katushka9-5 (37)	mRuby3 (24) mRuby3 Addgene (24)

Table 1: List of FPs Grouped By Common Ancestor. Citations (in parentheses) indicate original publication describing each variant.

REFERENCES

1. Balleza E, Kim JM, & Cluzel P (in press) Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods*.
2. Zaslaver A*, et al.* (2006) A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. *Nat Methods* 3(8):623-628.
3. Davis JH, Rubin AJ, & Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* 39(3):1131-1141.
4. Sugiura S, Ohkubo S, & Yamaguchi K (1993) Minimal essential origin of plasmid pSC101 replication: requirement of a region downstream of iterons. *J Bacteriol* 175(18):5993-6001.
5. Vocke C & Bastia D (1983) Primary structure of the essential replicon of the plasmid pSC101. *Proc Natl Acad Sci U S A* 80(21):6557-6561.
6. Cox RS, 3rd, Dunlop MJ, & Elowitz MB (2010) A synthetic three-color scaffold for monitoring genetic regulation and noise. *J Biol Eng* 4:10.
7. Tsien RY (1998) The green fluorescent protein. *Annu Rev Biochem* 67:509-544.
8. Nagai T*, et al.* (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol* 20(1):87-90.
9. Rekas A, Alattia JR, Nagai T, Miyawaki A, & Ikura M (2002) Crystal structure of venus, a yellow fluorescent protein with improved maturation and reduced environmental sensitivity. *J Biol Chem* 277(52):50573-50578.
10. Choi PJ, Cai L, Frieda K, & Xie S (2008) A stochastic single-molecule event triggers phenotype switching of a bacterial cell. *Science* 322(5900):442-446.
11. Kremers GJ, Goedhart J, van Munster EB, & Gadella TW, Jr. (2006) Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius. *Biochemistry* 45(21):6570-6580.
12. Markwardt ML*, et al.* (2011) An improved cerulean fluorescent protein with enhanced brightness and reduced reversible photoswitching. *PLoS One* 6(3):e17896.
13. Rizzo MA, Springer GH, Granada B, & Piston DW (2004) An improved cyan fluorescent protein variant useful for FRET. *Nat Biotechnol* 22(4):445-449.
14. Goedhart J*, et al.* (2010) Bright cyan fluorescent protein variants identified by fluorescence lifetime screening. *Nat Methods* 7(2):137-139.
15. Goedhart J*, et al.* (2012) Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%. *Nat Commun* 3.
16. Costantini LM*, et al.* (2015) A palette of fluorescent proteins optimized for diverse cellular environments. *Nat Commun* 6:7670.
17. Cubitt AB, Woollenweber LA, & Heim R (1999) Understanding structure-function relationships in the Aequorea victoria green fluorescent protein. *Methods Cell Biol* 58:19-30.
18. Zapata-Hommer O & Griesbeck O (2003) Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP. *BMC Biotechnol* 3:5.
19. Cormack BP, Valdivia RH, & Falkow S (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* 173(1):33-38.
20. Shaner NC, Steinbach PA, & Tsien RY (2005) A guide to choosing fluorescent proteins. *Nat Methods* 2(12):905-909.
21. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, & Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 24(1):79-88.
22. Nguyen AW & Daugherty PS (2005) Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat Biotechnol* 23(3):355-360.
23. Lam AJ*, et al.* (2012) Improving FRET dynamic range with bright green and red fluorescent proteins. *Nat Methods* 9(10):1005-1012.

24. Bajar BT*, et al.* (2016) Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Sci Rep* 6:20889.
25. Bevis BJ & Glick BS (2002) Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed). *Nat Biotechnol* 20(1):83-87.
26. Strack RL*, et al.* (2008) A noncytotoxic DsRed variant for whole-cell labeling. *Nat Methods* 5(11):955-957.
27. Campbell RE*, et al.* (2002) A monomeric red fluorescent protein. *Proc Natl Acad Sci U S A* 99(12):7877-7882.
28. Bindels DS*, et al.* (2017) mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat Methods* 14(1):53-56.
29. Shaner NC*, et al.* (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp red fluorescent protein. *Nat Biotechnol* 22(12):1567-1572.
30. Shen Y, Chen Y, Wu J, Shaner NC, & Campbell RE (2017) Engineering of mCherry variants with long Stokes shift, red-shifted fluorescence, and low cytotoxicity. *PLoS One* 12(2):e0171257.
31. Strack RL*, et al.* (2009) A Rapidly Maturing Far-Red Derivative of DsRed-Express2 for Whole-Cell Labeling (vol 48, pg 8279, 2009). *Biochemistry* 48(40):9704-9704.
32. Merzlyak EM*, et al.* (2007) Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nat Methods* 4(7):555-557.
33. Shaner NC*, et al.* (2008) Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat Methods* 5(6):545-551.
34. Shcherbo D*, et al.* (2009) Far-red fluorescent tags for protein imaging in living tissues. *Biochem J* 418(3):567-574.
35. Chu J*, et al.* (2014) Non-invasive intravital imaging of cellular differentiation with a bright red-excitable fluorescent protein. *Nat Methods* 11(5):572-578.
36. Shcherbo D*, et al.* (2007) Bright far-red fluorescent protein for whole-body imaging. *Nat Methods* 4(9):741-746.
37. Shcherbo D*, et al.* (2010) Near-infrared fluorescent proteins. *Nat Methods* 7(10):827-U1520.