

## Nature Methods

# Molecular evolution approaches to design advanced red fluorescent proteins

Fedor V Subach, Kiryl D Piatkevich & Vladislav V Verkhusha

**Supplementary Figure 1.** Alignment of amino acid sequences of derivatives of RFPs with different phenotype.

**Supplementary Table 1.** Overview of suboptimal properties of current monomeric red fluorescent proteins.

**Supplementary Table 2.** Current red fluorescent proteins exhibiting the best specific property or maximally optimized for the specific imaging application.

**Supplementary Table 3.** Properties of the major chromophore structures observed in fluorescent proteins.

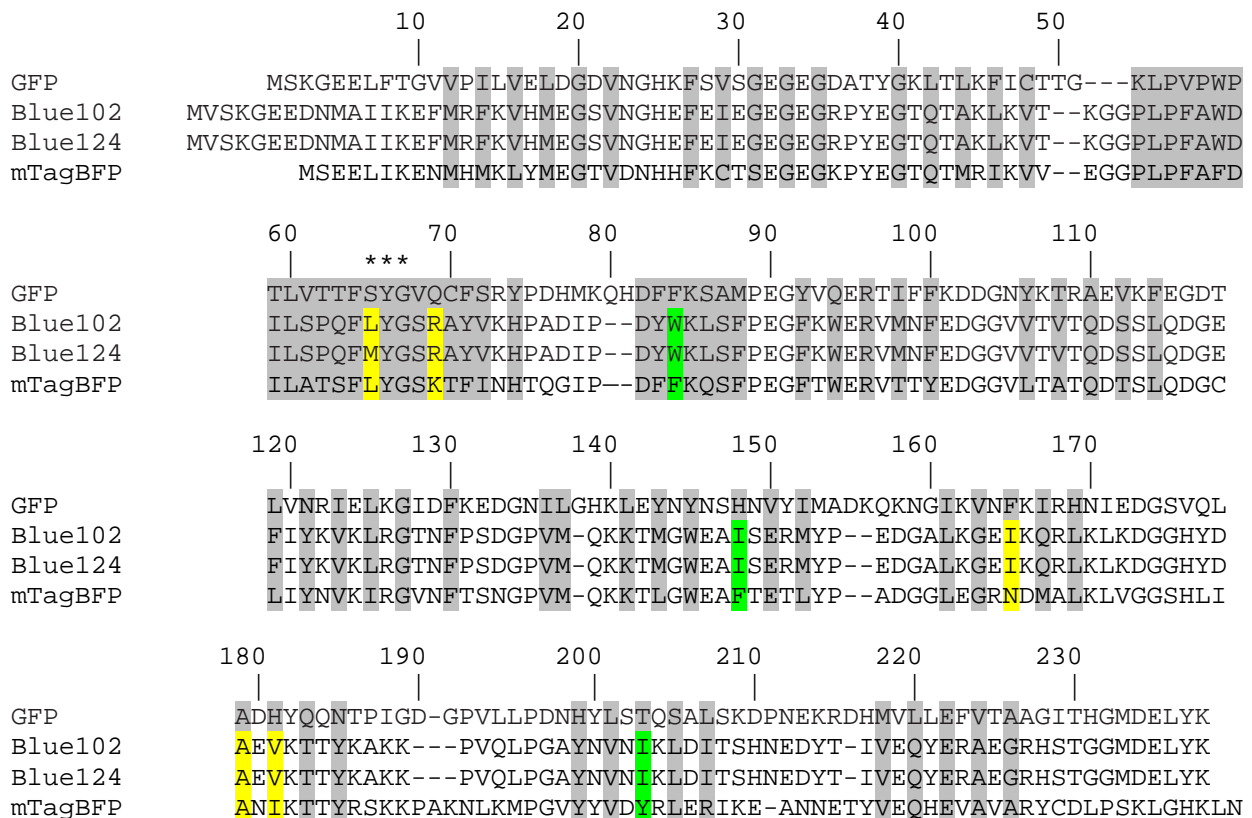
**Supplementary Table 4.** Amino acid residues in the supporting positions responsible for different phenotypes and properties of red fluorescent proteins.

**Supplementary Note**

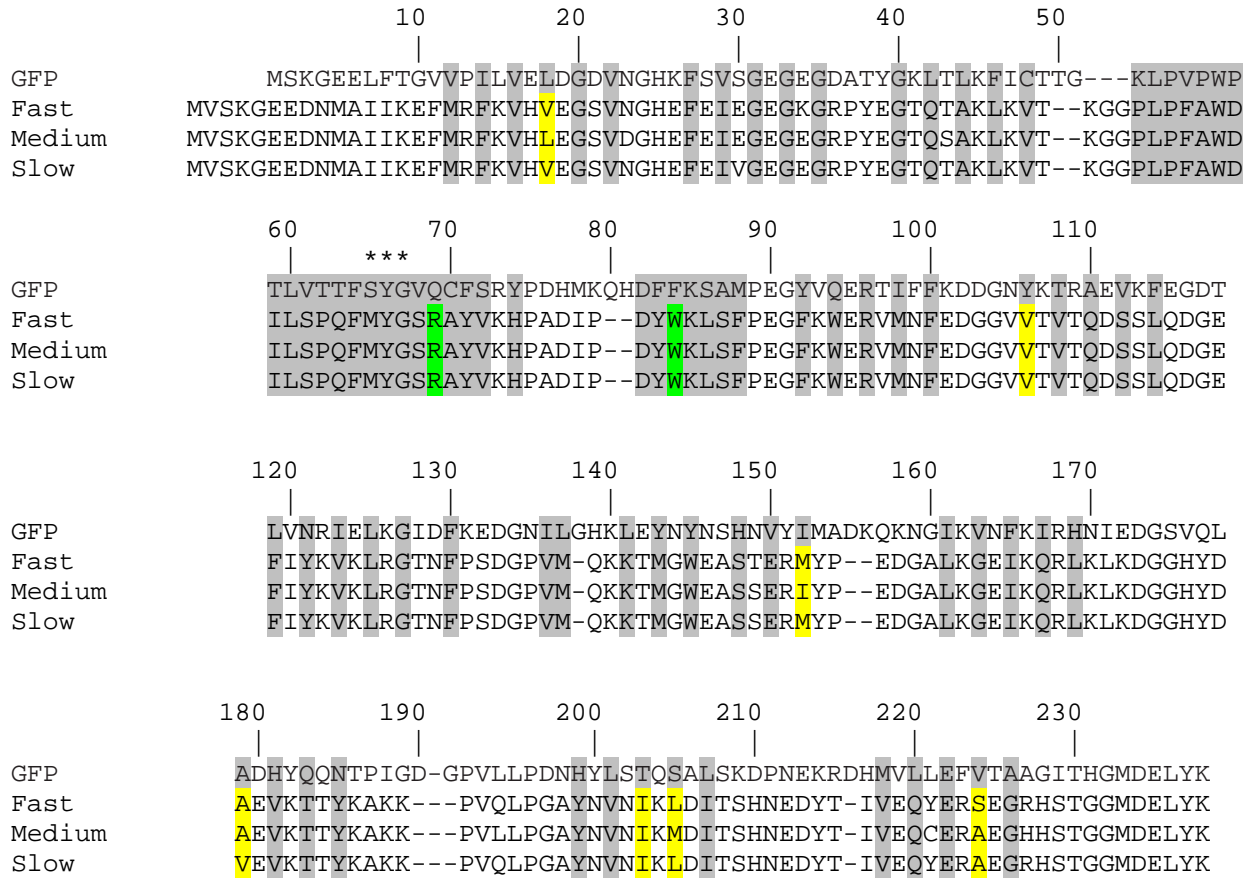
**Supplementary Figure 1.** Alignment of amino acid sequences of derivatives of RFPs with different phenotype.

(A) TagBFP-like phenotype, (B) Fluorescent timer-like phenotype, (C) Photoactivatable-like phenotype, (D) Photoswitchable-like phenotype, (E) Far-red-shifted-like phenotype, (F) Large Stokes shift-like phenotype; and properties: (G) High quantum yield, (H) High photostability, (I) High pH-stability, (J) Fast maturation, (K) Monomeric state. Residues in key positions are selected with red, green, or blue colors to denote which of them are responsible for the respective phenotype independently, in concerted manner, or either of these two ways, respectively. Amino acid residues in the supporting positions are highlighted with yellow color.

**(A) TagBFP-like phenotype**



**(B) Fluorescent timer-like phenotype**



### (C) Photoactivatable-like phenotype

```

          10      20      30      40      50
          |      |      |      |      |
GFP      MSKGEELFTGVVPIILVELDGDVNGHKFSVSGEGEGDATYGKLTTLKFICTTG---KLPVPWP
PAmCherry1 MVSKEEDNMAIIKEFMRFKVHMEGTVNGHVFEIEGEGEGRPYEGTQTAKLKVT--KGGPLPFTWD
PAmCherry2 MVSKEEDNMAIIKEFMRFKVHLEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT--KGGPLPFAWD
PAmCherry3 MVSKEEDNMAIIKEFMRFKVHLEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT--KGGPLPFTWD
PATagRFP      MSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD
PAmKate      MSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD

          60      70      80      90      100     110
          |      |      |      |      |      |
GFP      TLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDT
PAmCherry1 ILSPQFMYGSNAYVKHPADIP--DYFKLSFPEGFKWERVMKFEDGGVVTVTQDSSLQDGE
PAmCherry2 ILSPQFMYGSNAYVKHPADIP--DYFKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
PAmCherry3 ILSPQFMYGSNAYVKHPADIP--DYFKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
PATagRFP      ILATSFMYGSTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
PAmKate      ILATSFMYGSKTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC

          120     130     140     150     160     170
          |      |      |      |      |      |
GFP      LVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQL
PAmCherry1 FIYKVKLRGTNFPDGPVM-QKKTMGWEALSERMYP--EDGALKGEVVKPRVKKLDGGHYD
PAmCherry2 FIYKVKLRGTNFPDGPVM-QKKTMGWETLSERMYP--EDGALKGELKARTKLDGGHYD
PAmCherry3 FIYKVKLRGTNFPDGPVI-QKKTMGWDALSERMYP--EDGALKGELKARLKLKLDGGHYE
PATagRFP      LIYNVKIRGVNFPDGPVM-KKKTLGWEPSTEKLKP--ADGGLEGRVDMALKLVGGGHLI
PAmKate      LIYNVKIRGVNFPDGPVM-QKKTLGWEANTEMLYP--ADGGLEGRGDMALKLVGGGHLI

          180     190     200     210     220     230
          |      |      |      |      |      |
GFP      ADHYQONTPIGD-GPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK
PAmCherry1 AEVKTTYKAKK---PVQLPGAYNVNRKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
PAmCherry2 TEVKTTYKAKK---PVQLPGAYNVNRKLDITSHNEDYT-IVEQYERAEGRLHSTGGMDELYK
PAmCherry3 AEVKTTYKAKK---PVQLPGAYNVNRKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
PATagRFP      CNFKTTYRSKKPAKNLKMFGVYVDRRLERIKEADKET-YWEQHEVAVARYSDLPSKLGHLN
PAmKate      CNLKTTYRSKKPAKNLKMFGVYVDRRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHLN
```

(D) Photoswitchable-like phenotype

```

          10      20      30      40      50
          |      |      |      |      |
GFP      MSKGEELFTGVVPI LVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTG---KLPVPWP
rsCherryRev  MVSKGEEDNMAIIKEFMRFKVHMEG SVNGHEFEIEEGEGEGRPYEGTQTAKLKVT--KGGPLPFAWD
rsCherry    MVSKGEEDNMAIIKEFMRFKVHMEG SVNGHEFEIEEGEGEGRPYEGTQTAKLKVT--KGGPLPFAWD
rsTagRFP    MVSKGEELIKENMHMKLYMEGT VNNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD
KFPl      MASLLTETMPFKTTIEGTVNGHCFKCIGKGEGNPFEGTQEMKIEVI--EGGPLPFAFH
IrisFP     HHMSAIKPD MKINLRMEGNVNGHHFVIDGDGTGKPFEGKQSM DLEVK--EGGPLPFAFD

          60      70      80      90      100     110
          |      |      |      |      |      |
GFP      TLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDT
rsCherryRev  ILSPQFMYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
rsCherry    ILSPQFMYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
rsTagRFP    ILATSFMYGSRTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
KFPl      ILSTSCMYGSKTFIKYVSGIP--DYFKQSFPEGFTWERTTTYEDGGFLT AHQDTSLDGDC
IrisFP     ILTTAFHYGNRVFAEYPDHIQ--DYFKQSFPGKYSWERSLTFEDGGIC IARNDITMEGDT

          120     130     140     150     160     170
          |      |      |      |      |      |
GFP      LVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIMADKQKNGIKVNFKIRHNI EDGSVQL
rsCherryRev  FIYKVKLRG TNFPSDGPVM-QKKTMGWVACSERMYP--EDGALKGESK MRLKLDGGHYD
rsCherry    FIYKVKLRG TNFPSDGPVM-QKKTMGWVASSERMYP--EDGALKGESK QRLKLDGGHYD
rsTagRFP    LIYNVKLRG VNFPSNGPVM-QKKT LGWEAATEMLYP--ADGGLEGRGDMALKLVGGGHLI
KFPl      LVYKVKILGN NFPADGPVM-QNKVGRWEPGTEIVYE--VDGVLRGQSLMALKCPGGRHLT
IrisFP     FYNKVRFHG VNF PANGPVM-QKKT LKWEPESTEKMYV--RDGVLTGDI TMALLLEGNAHYR

          180     190     200     210     220     230
          |      |      |      |      |      |
GFP      ADHYQQNTPIGD-GPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK
rsCherryRev  AEFKTTYKAKK---PVQLPGAYNVN IKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
rsCherry    AEFWTTYKAKK---PVQLPGAYNVN IKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
rsTagRFP    CNLKT TYRSKNPAK NLKMPGVYFVDHRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHKL N
KFPl      CHLHT TYRSK KPASAL KMPGFHFEDHRIEIMEEVEK GK-CYKQYEA AVGRYCD AAPS KLGHN
IrisFP     CDSRT TYKAKE--KGVKLPGYHLVDHCIEILSHDKDYN-KVKLYEHAVAH SGLPDNARR
```

### (E) Far-red-shifted-like phenotype

```

          10      20      30      40      50
GFP      MSKGEELFTGVVPIILVELDGDVNGHKFSVSGEGEGDATYGKLTTLKFICTTG---KLPVPWP
mNeptune MSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTGRIKVV--EGGPLPFAFD
TagRFP657 MSELITENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTORIKVV--EGGPLPFAFD
E2-Crimson MDSTENVIKPFMRFKVHMEGSVNGHEFEIEGVGEGKPYEGTQTAKLQVT--KGGPLPFAWD
eqFP650   MGEDSELISENMHMKLYMEGTVNGHHFKCTSEGEGKPYEGTQTAKIKVV--EGGPLPFAFD
eqFP670   MGEDSELISENMHTKLYMEGTVNGHHFKCTSEGEGKPYEGTQTCKIKVV--EGGPLPFAFD
mPlum     MVSKGEENMAIIKEFMRFKEHMEGSVNGHEFEIEEGEGGRPYEGTQTARLKVT--KGGPLPFAWD
mRouge    MVSKGEEDNMAIIKEFMRFKTHMEGSVNGHEFEIEEGEGGRPYEGTQTAKLKVT--KGGPLPFAWD
RFP639    MNSLIKENMRMMVMMEGSVNGYQFKCTGEGDGNPYMGTQTRIKVV--EGGPLPFAFD

          60      70      80      90      100     110
GFP      TLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDT
mNeptune ILATCFMYGSKTFINHTQGIP--DFFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
TagRFP657 ILATSFMYGSHTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
E2-Crimson ILSPQFFYGSKAYIKHPADIP--DYLKQSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGT
eqFP650   ILATSFMYGSKTFINHTQGIP--DFFKQSFPEGFTWERITTYEDGGVLTATQDTSLQNGC
eqFP670   ILATSFMYGSKTFINHTQGIP--DFFKQSFPEGFTWERITTYEDGGVLTATQDTSLQNGC
mPlum     ILSPQIQYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
mRouge    ILSPQFMYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
RFP639    ILATSFMYGSKTFIKHTKGIP--DFFKQSFPEGFTWERVTRYEDGGVFTVMQDTSLQEDGC

          120     130     140     150     160     170
GFP      LVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQL
mNeptune LIYNVKIRGVNFPSNGPVM-QKKTLGWEASTETLYP--ADGGLEGRCDMALKLVGGGHLI
TagRFP657 LIYNVKIRGVNFPSNGPVM-QKKTLGWEAHTETLYP--ADGGLEGRTALAKLVGGGHLI
E2-Crimson LIYHVKFIGVNFPSDGPVM-QKKTLGWEPSTERNYP--RDGVLKGENHMALKLVGGGHYL
eqFP650   LIYNVKINGVNFPSNGPVM-QKKTLGWEASTETLYP--ADSGLRGHSQMALKLVGGGYLH
eqFP670   LIYNVKINGVNFPSNGPVM-QKKTLGWEANTETLYP--ADSGLRGHNQMALKLVGGGYLH
mPlum     FIYKVKVRGTNFPSDGPVM-QKKTMGWEASTERMYP--EDGALKGEMKMRLRLKLDGGGHYD
mRouge    FIYKVKLRGTNFPSDGPVM-QKKTMGWEACSERMYP--EDGALKGEMKMRLRLKLDGGGHYD
RFP639    LVYHAKVTGVNFPSNGAVM-QKKTKGWEPSTETLYP--ADGGLRGYCQMALNVDGGGYLF

          180     190     200     210     220     230
GFP      ADHYQQNTPIGD-GPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYK
mNeptune CNLKTTYRSKKPAKNLKMPGVYFVDRRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHKLN
TagRFP657 CNFKTTYRSKKPAKNLKMPGVYYVDYRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHKLN
E2-Crimson CEFKSIYMAKPP---VKLPGYHYVDYKLDITSHNEDYT-VVEQYERAEARHHLFQ
eqFP650   CSLKTTYRSKKPAKNLKMPGFYFVDRKLERIKEADKET-YVEQHEMAVARYCDLPSKLGHS
eqFP670   CSLKTTYRSKKPAKNLKMPGFYFVDRKLERIKEADKET-YVEQHEMAVARYCDLPSKLGHS
mPlum     AEVKTTYMAKPP---VQLPGAYKTDIKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
mRouge    AEVKTTYKAKPP---VQLPGAYNTNYKLDITSHNEDYT-IVEQYERNEGRHSTGGMDELYK
RFP639    CSFETTYRSKKTDENFKMPGFFHVDHRLERLEESDKEM-FVVQHEHAVAKFCDLPSKLGRL
```

**(F) Large Stokes shift-like phenotype**

	10	20	30	40	50
GFP	MSKGEELFTGVVPI	LVELDGDVNGHKFS	SVSGEGEGDATY	GKLTTLKFICTTG	---KLPVPWP
LSSmKate1	MSELITENMHMKLY	MEGTVNNHHFKCT	SEGEKPYEGTQT	MRKVV--EGG	PLPFAFD
LSSmKate2	MSELITENMHMKLY	MEGTVNNHHFKCT	SEGEKPYEGTQT	MRKVV--EGG	PLPFAFD
mKate/158G/160E	MSELITENMHMKLY	MEGTVNNHHFKCT	SEGEKPYEGTQT	MRKVV--EGG	PLPFAFD
mKate/143D/158V/160S	MSELIKENMHMKLY	MEGTVNNHHFKCT	SEGEKPYEGTQT	MRKVV--EGG	PLPFAFD
mNeptune/158D	MSELIKENMHMKLY	MEGTVNNHHFKCT	SEGEKPYEGTQT	GRIKVV--EGG	PLPFAFD
mCherry/158E/160A	SKGEEDNMAIIKEF	MRFKVHMEGSVNG	HEFEIEEGEGEGR	PYEGTQTAKL	KVT--KGG
mStrawberry/158D	VSKGEENNMAIIKEF	MRFKVRMEGSVNG	HEFEIEEGEGEGR	PYEGTQTAKL	KVT--KGG
mOrange/158D/160G	SKGEENNMAIIKEF	MRFKVRMEGSVNG	HEFEIEEGEGEGR	PYEGFQTAKL	KVT--KGG
mKO/158E	MVSVIKPEMKMRY	YMDGSVNGHEFT	IEEGEGTGRPYE	GHQEMTLRV	TMAKGGPMP
mKeima	MVSVIAKQMTYK	VYMSGTVNGHYF	EVEGDGKPKPYE	GEQTVKLT	TVT--KGG
	60	70	80	90	100
GFP	TLVTTFSYGVQCF	SRYPDHMKQHDF	FKSAMPEGYVQ	ERTIFFKDDG	NYKTRA
LSSmKate1	ILATSFMYGSYTF	INHTQGIP--DFF	KQSFPEGFTW	ERVTTYEDG	GVLTATQ
LSSmKate2	ILATSFMYGSYTF	INHTQGIP--DFF	KQSFPEGFTW	ERVTTYEDG	GVLTATQ
mKate/158G/160E	ILATSFMYGSKTF	INHTQGIP--DFF	KQSFPEGFTW	ERVTTYEDG	GVLTATQ
mKate/143D/158V/160S	ILATSFMYGSR	TFINHTQGIP--DFF	KQSFPEGFTW	ERVTTYEDG	GVLTATQ
mNeptune/158D	ILATCFMYGSK	TFINHTQGIP--DFF	KQSFPEGFTW	ERVTTYEDG	GVLTATQ
mCherry/158E/160A	ILSPQFMYGSK	AYVKHPADIP--DYL	KLSFPEGFKW	ERVMNFEDG	GVVTVTQ
mStrawberry/158D	ILTPNFYGSK	AYVKHPADIP--DYL	KLSFPEGFKW	ERVMNFEDG	GVVTVTQ
mOrange/158D/160G	ILSPQFTYGS	KAYVKHPADIP--DYF	KLSFPEGFKW	ERVMNFEDG	GVVTVTQ
mKO/158E	LVSHVFCYGHR	PFTKYPEEIP--DYF	KQAFPEGLS	WERSLEFED	GGSASVSAH
mKeima	ILSPQLQYGS	IPFTKYPEDIP--DYF	KQSFPEGYT	WERSMNFED	GAVCTVSND
	120	130	140	150	160
GFP	LVNRIELKGI	DFKEDGNILGH	KLEYNYN	SHNVYIMADK	QKNGIKV
LSSmKate1	LIYNVKIRGV	NFTSNGPVM-Q	KKTLGWEAG	TEMLYP--AD	GGLEGRS
LSSmKate2	LIYNVKIRGV	NFTSNGPVM-Q	KKTLGWEAG	TEMLYP--AD	GGLEGRS
mKate/158G/160E	LIYNVKIRGV	NFP	SNPVM-QKKT	LGWEASTE	MLYP--AD
mKate/143D/158V/160S	LIYNVKIRGV	NFP	SNPVM-QKKT	LGWEADTE	MLYP--AD
mNeptune/158D	LIYNVKIRGV	NFP	SNPVM-QKKT	LGWEASTE	TETLYP--AD
mCherry/158E/160A	FIYKVKLRGT	NFP	PSDGPVM-Q	KKTMGWEA	SSEMYP--ED
mStrawberry/158D	FIYKVKLRGT	NFP	PSDGPVM-Q	KKTMGWEA	SSEMYP--ED
mOrange/158D/160G	FIYKVKLRGT	NFP	PSDGPVM-Q	KKTMGWEA	SSEMYP--ED
mKO/158E	FYHKSFTG	VNFPADGP	IM-QNQSV	DWEPSTEK	ITA--SDG
mKeima	FIYNVKIS	GENFPPNGP	VIM-QKKT	QGWEPSTE	RLFA--RD
	180	190	200	210	220
GFP	DHYQQNTPI	GD-GPVLL	PDNHYLSTQ	SALS	KDPNEKRD
LSSmKate1	NLKSTYRSK	PKAKNLKVP	GVYVDRRL	LERIKEAD	KET-YVE
LSSmKate2	NLKSTYRSK	PKAKNLKVP	GVYVDRRL	LERIKEAD	KET-YVE
mKate/158G/160	NLKTTYRSK	PKAKNLKMP	GVYVDRRL	LERIKEAD	KET-YVE
mKate/143D/158V/160S	NLKTTYRSK	PKAKNLKMP	GVYVDRRL	LERIKEAD	KET-YVE
mNeptune/158D	NLKTTYRSK	PKAKNLKMP	GVYVDRRL	LERIKEAD	KET-YVE
mCherry/158E/160A	EVKTTYK	AKKPV---QL	P	GAYVNIK	L
mStrawberry/158D	EVKTTYK	AKKPV---QL	P	GAYVNIK	L
mOrange/158D/160G	EVKTTYK	AKKPV---QL	P	GAYVNIK	L
mKO/158E	QFKTTYK	AAKKI--L	KMPGSHYI	SHRLVRK	TEGN---
mKeima	EFKTTYK	AKKPV---R	MPGRHEI	DRKLDVT	SHNRD
	230				

### (G) High quantum yield

10 20 30 40 50

GFP MSKGEELFTGVVPIILVELDGDVNGHFKFSVSGEGEGDATYGKLTLLKFICTTG--KLPVPWP  
rsTagRFP MVSKGEELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD  
KFP1 MASLLTETMPFKTTIEGTVNGHCFKCIKKGEGNPFEGTQEMKIEVI--EGGPLPFAFH  
asFP595\_70A/148S MASFLKKTMPFKTTIEGTVNGHYFKCTGKKGEGNPFEGTQEMKIEVI--EGGPLPFAFH  
HcRed1 MVSGLLKESMRIKMYMEGTVNGHYFKCEGEGDGNPFAGTQSMRIHVT--EGAPLPFAFD  
mCherry MVSKGEEDNMAIIKEFMRFKVHMEGTVNGHEFEIEGEGEGRPYEGTQAKLKVT--KGGPLPFAWD  
mKate2 MVSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKAV--EGGPLPFAFD  
mOrange MVSKGEENMAIIKEFMRFKVRMEGTVNGHEFEIEGEGEGRPYEGFQAKLKVT--KGGPLPFAWD  
TagRFP MVSKGEELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD

60 70 80 90 100 110

GFP TLVTTFSYGVQCFSTRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEFEGDT  
rsTagRFP ILATSFMYGSRTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC  
KFP1 ILSTSCMYGSKTFIKYVSGIP--DYFKQSFPEGFTWERTTTYEDGGFLTAHQDTSLDGDC  
asFP595\_70A/148S ILSTSCMYGSKAFIKYVSGIP--DYFKQSFPEGFTWERTTTYEDGGFLTAHQDTSLDGDC  
HcRed1 ILAPCCEYGSRTFVHHTAEIP--DFFKQSFPEGFTWERTTTYEDGGILTAAHQDTSLEGN  
mCherry ILSPQFMYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVLTATQDSSSLQDGE  
mKate2 ILATSFMYGSKTFINHTQGIP--DFFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC  
mOrange ILSPQFTYGSKAYVKHPADIP--DYFKLSFPEGFKWERVMNFEDGGVVTATQDSSSLQDGE  
TagRFP ILATSFMYGSRTFINHTQGIP--DFFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC

120 130 140 150 160 170

GFP LVNRIELKGIDFKEDGNILGHKLEYNNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQL  
rsTagRFP LIYNVKLRGVNFPNSNGPVM-QKKTLGWEAAATEMLYP--ADGGLEGRGDMALKLVGGGHLI  
KFP1 LVYKVKILGNNFPADGPVM-QNKVGRWEPGTEIVYE--VDGVLRGQSLMALKKCPGGRHLT  
asFP595\_70A/148S LVYKVKILGNNFPADGPVM-QNKAGRWEPESTEIVYE--VDGVLRGQSLMALKKCPGGRHLT  
HcRed1 LIYKVKVLGTNFPADGPVM-KNKSGGWEPESTEIVYE--ENGVLGRNVMALKV-GDRHLI  
mCherry FIYKVKLRGTNFPSDGPVM-QKKTMGWEASSERMYP--EDGALKGEIKQRLKLDGGHYD  
mKate2 LIYNVKIRGVNFPNSNGPVM-QKKTLGWEASTETLYP--ADGGLEGRADMALKLVGGGHLI  
mOrange FIYKVKLRGTNFPSDGPVM-QKKTMGWEASSERMYP--EDGALKGEIKMRLKLDGGHYT  
TagRFP LIYNVKIRGVNFPNSNGPVM-QKKTLGWEANTEMLYP--ADGGLEGRSDMALKLVGGGHLI

180 190 200 210 220 230

GFP ADHYQQNTPIGD-GPVLLPDNHYLSTQSALS KDPNEKRDMVLLLEFVTAAGITHGMDELYK  
rsTagRFP CNLKTTYRSKNPAKNLKMPGVYFVDHRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHLN  
KFP1 CHLHTTYRSKPPASALKMPGFHFEDHRIEIMEEVEKKGK-CYKQYEA AVGRYCDAAAPSKLGHN  
asFP595\_70A/148S CHLHTTYRSKPPASALKMPGFHFEDHRIEIMEEVEKKGK-CYKQYEA AVGRYCDAAAPSKLGHN  
HcRed1 CHHYTSYRSKKA VRALTMPGFHFDTIRLQMLRKEKDE--YFELYEASVARYSDLPEKAN  
mCherry AEVKTTYKAKK---PVQLPGAYNVN IKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK  
mKate2 CNLKTTYRSKPPAKNLKMPGVYVDRRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHLN  
mOrange SEVKTTYKAKK---PVQLPGAYIVG IKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK  
TagRFP CNFKTTYRSKPPAKNLKMPGVYVDRRLERIKE-ADKETYVEQHEVAVARYCDLPSKLGHLN



## (H) High photostability

10 20 30 40 50

GFP MSKGEELFTGVVPIILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTG---KLPVPWP  
TagRFP-T MVSKGEELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD  
mOrange2 MVSKGEENMAIIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGFQTAKLKVT--KGGPLPFAWD  
mKate2 MVSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRIKAV--EGGPLPFAFD

60 70 80 90 100 110

GFP TLVTTFSYGVQCFSTRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDT  
TagRFP-T ILATSFMYGSRTFINHTQGIP--DFFKQSFPEGFTWERVTTYEDGGVLTATQDTS LQDGC  
mOrange2 ILSPHFTYGSKAYVKHPADIP--DYFKLSFPEGFKWERVMNYEDGGVVTVTQDSSLQDGE  
mKate2 ILATSFMYGSKTFINHTQGIP--DFFKQSFPEGFTWERVTTYEDGGVLTATQDTS LQDGC

120 130 140 150 160 170

GFP LVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQL  
TagRFP-T LIYNVKIRGVNFP SNGPVM-QKKT LGWEANTEMLYP--ADGGLEGRITDMALKLVGGGHLI  
mOrange2 FIYKVKLRGTNFP SNGPVM-QKKT MGWEASSERMYPE--DGALKGKI KMRLKLDGGHYT  
mKate2 LIYNVKIRGVNFP SNGPVM-QKKT LGWEASTETLYPA--DGGLEGRAD MALKLVGGGHLI

180 190 200 210 220 230

GFP ADHYQQNTPIGD-GPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK  
TagRFP-T CNFKTTYRSKPKAKNLKMPGVYVDRRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHKLN  
mOrange2 SEVKTTYKAKK---PVQLPGAYIVDIKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK  
mKate2 CNLKTTYRSKPKAKNLKMPGVYVDRRLERIKEADKET-YVEQHEVAVARYCDLPSKLGHKLN

## (I) High pH-stability

10 20 30 40 50

GFP MSKGEELFTGVVPI LVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTG---KLPVPWP  
mTagBFP MSEELIKENMHMKLYMEGTVDNHHFKCTSEGEGKPYEGTQTMRIKVV--EGGPLPFAFD  
DsRed2 MASSENVITEFMRFKVRMEGTVNGHEFEIEGEGEGRPHYEGHNTVKKLKV---KGGPLPFAWD

60 70 80 90 100 110

GFP TLVTTFSYGVQCF SRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDT  
mTagBFP ILATSFLYGSKTFINHTQGIP--DFFKQSFPEGFTWERVTTTYEDGGVLTATQDTS LQDGC  
DsRed2 ILSPQFQYGSKVYVKHPADIP--DYKKLSFPEGFKWERVMNFEDGGVATVTQDSSLQDGC

120 130 140 150 160 170

GFP LVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQL  
mTagBFP LIYNVKIRGVNFTSNGPVM-QKKT LGWEAFTETLYP--ADGGLEGRNDMALKLVGGSHLI  
DsRed2 FIYKVKFIGVNFPSDGPVM-QKKT MGWEASTERLYP--RDGVLKGETHKALKLKDGGHYL

180 190 200 210 220 230

GFP ADHYQQNTPIGD-GPVLLPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK  
mTagBFP ANIKTTYRSKKPAKNLKM PGVYYVDYRLERIKEANNET-YVEQHEVAVARYCDLPSKLGHKLN  
DsRed2 VEFKSIYMAKPP---VQLPGY YVDAKLDITSHNEDYT-IVEQYERTEGRHHLFL

(J) Fast maturation

```

          10      20      30      40      50
GFP      MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTTLKFICTTG---KLPVPWP
mKate2   MVSELIKENMHMKLYMEGTVNNHHFKCTSEGEGKPYEGTQTMRKAV--EGGPLPFAFD
mCherry  MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVT--KGGPLPFAWD
SlowFT   MVSKGEEDNMAIIKEFMRFKVHVEGSVNGHEFEIVGEGEGRPYEGTQTAKLKVT--KGGPLPFAWD
mRuby    MNSLIKENMRMKVVLSEGSVNGHQFKCTGEGEGNPYMGTTMRKVI--EGGPLPFAFD

          60      70      80      90      100     110
GFP      TLVTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEEDT
mKate2   ILATSFMYGSKTFINHTQGIP--DFFKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
mCherry  ILSPQFMYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
SlowFT   ILSPQFMYGSRAYVKHPADIP--DYWKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
mRuby    ILATSFMYGSRTFIKYKPGIP--DFFKQSFPEGFTWERVTRYEDGGVITVMQDTSLEDGC

          120     130     140     150     160     170
GFP      LVNRIELKGIDFKEDGNILGHKLEYNYNshNVYIMADKQKNGIKVNFKIRHNIEDGSVQL
mKate2   LIYNVKIRGVNFPSNGPVM-QKKTLGWEASTETLYPA--DGGLEGRADMALKLVGGGHLI
mCherry  FIYKVKLRGTNFPDGPVM-QKKTMGWEASSERMYP--EDGALKGEIKQRLKLDGGHYD
Slow     FIYKVKLRGTNFPDGPVM-QKKTMGWEASSERMYP--EDGALKGEIKQRLKLDGGHYD
mRuby    LVYHAQVRGVNFPSNGAVM-QKKTGWEPNTEMYP--ADGGLRGYTHMALKVDGGGHLs

          180     190     200     210     220     230
GFP      ADHYQQNTPIGD-GPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTAAGITHGMDELYK
mKate2   CNLKTTYRSKKPAKNLKMPGVYYVDRRLERIKEADKET-YVEQHEVAVARYCDLPSKLGKLN
mCherry  AEVKTTYKAKK---PVQLPGAYNVNIKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
SlowFT   VEVKTTYKAKK---PVQLPGAYNVNIKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
mRuby    CSFVTTYRSKKTVGNIKMPGIHAVDHRLELERLEESDNEM-FVVQREHAVAKFAGLGGGSLRSR
```

### (K) Monomeric state

```

          10      20      30      40      50
GFP      MSKGEELFTGVVPIILVELDGDVNGHKFSVSGEGEGDATYGKLTLLKFICTTG---KLPVWPW
PATagRFP MSELIKENMHMKLYMEGTVNNHHFKCTSEGEKPYEGTQTMRIKVV--EGGPLPFAFD
mNeptune MSELIKENMHMKLYMEGTVNNHHFKCTSEGEKPYEGTQTRIKVV--EGGPLPFAFD
mCherry  MVSKEEDNMAIIKEFMRFKVMHEGSVNGHEFEIEGEGEGRPYEGTQAKLKVT--KGGPLPFAWD
mRuby    MNSLIKENMRMKVVLEGSVNGHQFKCTGEGEGNPYMGQTQTMRIKVI--EGGPLPFAFD

          60      70      80      90      100     110
GFP      TLVTTFSYGVQCFSTRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEEDT
PATagRFP ILATSFMYGSSTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
mNeptune ILATCFMYGSKTFINHTQGIP--DFWKQSFPEGFTWERVTTYEDGGVLTATQDTSLQDGC
mCherry  ILSPQFMYGSKAYVKHPADIP--DYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGE
mRuby    ILATSFMYGSRTFIKYKPGIP--DFWKQSFPEGFTWERVTRYEDGGVITVMQDTSLEDGC

          120     130     140     150     160     170
GFP      LVNRIELKGIIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHNIEDGSVQL
PATagRFP LIYNVKIRGVNFPSNGPVM-KKKTLGWEPSTEKLP--ADGGLEGRVDMALKLVGGGHLI
mNeptune LIYNVKIRGVNFPSNGPVM-QKKTLGWEASTETLYP--ADGGLEGRCDMALKLVGGGHLI
mCherry  FIYKVKLRGTNFPSDGPVM-QKKTMGWEASSERMYP--EDGALKGEIKQRLKLDGGHYD
mRuby    LVYHAQVRGVNFPSNGAVM-QKKTGWEPNTEMMYP--ADGGLRGYTHMALKVDGGGHLI

          180     190     200     210     220     230
GFP      ADHYQQNTPIGD-GPVLDPDNHYLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYK
PATagRFP CNFKTTYRSKPKAKNLKMPGVYVDRRLEIIEADKET-YWEQHEVAVARYSDLPSKLGHKLN
mNeptune CNLKTTYRSKPKAKNLKMPGVYFVDRRLERIEADKET-YVEQHEVAVARYCDLPSKLGHKLN
mCherry  AEVKTTYKAKK---PVQLPGAYNVNIKLDITSHNEDYT-IVEQYERAEGRHSTGGMDELYK
mRuby    CSFVTTYRSKKTVGNIKMPGIHAVDHLRLERLEESDNEM-FVVQREHAVAKFAGLGGGSGLSR
```

**Supplementary Table 1.** Overview of suboptimal properties of current monomeric red fluorescent proteins.

Phenotype of RFP	Fluorescent protein	Form and its Excitation/Emission maxima, nm	Suboptimal properties	Ref.
Fluorescent	mKO	548/559	maturation, photoconversion to red species	1
	mKO2	551/565	photostability	2
	mOrange	548/562	photostability, pH-stability, maturation, photoconversion to far-red species	3
	mOrange2	549/565	pH-stability, maturation, photoconversion to far-red species	4
	TagRFP	555/584	maturation, photostability, monomeric state	5
	TagRFP-T	555/584	maturation, monomeric state	4
	mRuby	558/605	maturation, photostability	6
	mStrawberry	574/596	quantum yield, photostability	3
	mRaspberry	574/596	maturation, quantum yield, photostability	7
	mCherry	587/610	quantum yield	3
	mKate2	588/633	pH-stability	8
	mPlum	590/649	maturation, brightness, photostability	7
	eqFP650	592/650	quantum yield, photostability	9
	mNeptune	600/650	quantum yield, maturation, monomeric state, photostability	10
eqFP670	605/670	quantum yield	9	
Large Stokes shift	mKeima	440/620	brightness, pH-stability, monomeric state, maturation, residual red fluorescence with yellow excitation light	11
	LSS-mKate1	463/624	brightness, maturation	12
	LSS-mKate2	460/605	brightness, maturation	
Fluorescent Timers	FastFT	Blue, 403/466	photostability	13
		Red, 583/606	quantum yield	
	MediumFT	Blue, 401/464	photostability	
		Red, 579/600	quantum yield	
	SlowFT	Blue, 402/583	photostability	

		Red, 465/604	quantum yield	
<b>Photoactivatable</b>	PAmCherry1	Dark, 404/466	N/A	14
		Red, 564/594	extinction coefficient, pH-stability	
	PATagRFP	Dark, 351/none	N/A	15
		Red, 562/595	maturation	
<b>Photoconvertible</b>	mEos2	Green, 506/519	photostability, maturation, oligomeric state	16
		Red, 573/584	pH stability	
	mKikGR	Green, 505/515	photostability, pH-stability	17
		Red, 580/591	photostability	
<b>Reversibly photoswitchable</b>	KFP1	Dark, 580/600	N/A	18
		Red, 580/600	maturation, quantum yield	
	rsTagRFP	Dark, 440/none	extinction coefficient	19
		Red, 567/585	brightness, pH stability, maturation, contrast	
	rsCherry	Dark, 572/610	contrast	14,20
		Red, 572/610	brightness, maturation, complex photobehavior during switching,	
	rsCherryRev	Dark, 572/608	contrast	
		Red, 572/608	brightness, maturation, photoactivation during switching,	

N/A – not applicable.

**Supplementary Table 2.** Current red fluorescent proteins exhibiting the best specific property or maximally optimized for the specific imaging application.

<b>RFP property</b>	<b>Conventional RFPs</b>	<b>Photoactivatable RFPs</b>	<b>RFP application</b>	<b>Conventional RFPs</b>	<b>Photoactivatable RFPs</b>
<i>Excitation beyond 610 nm</i>	TagRFP657 <sup>21</sup>	not available	<i>STED</i>	TagRFP657	not applicable
<i>Emission beyond 670 nm</i>	eqFP670 <sup>9</sup>	not available	<i>Deep tissue imaging</i>	mNeptune and E2-Crimson	not available
<i>One photon brightness</i>	mKOk <sup>22</sup> and tdTomato <sup>3</sup>	mEos2 <sup>16</sup>	<i>RESOLFT</i>	not applicable	rsTagRFP
<i>Two-photon brightness</i>	DsRed2 and dTomato <sup>23</sup>	unknown	<i>PALM</i>	not applicable	PATagRFP and tdEosFP
<i>Photon counts per molecule</i>	Unknown	tdEosFP <sup>24</sup>	<i>Multicolor PALM</i>	not applicable	green PA-GFP
<i>pH stability</i>	LSS-mKate2 <sup>12</sup> and TagRFP <sup>5</sup>	PATagRFP <sup>15</sup> and KikGR <sup>25</sup>			red PATagRFP
<i>Maturation rate</i>	mCherry <sup>3</sup> and mKate2 <sup>8</sup>	rsTagRFP <sup>19</sup>	<i>Multicolor imaging with a single excitation wavelength</i>	blue TagBFP	not available
<i>Intracellular half-life and thermodynamic stability</i>	mRFP1 <sup>26</sup>	Dendra2 <sup>27</sup>		green T-Sapphire	
<i>Cytotoxicity</i>	DsRed-Express2 <sup>28</sup> and E2-Crimson <sup>29</sup>	unknown		red LSS-mKate2	
<i>Monomeric state and performance in fusions</i>	mCherry <sup>3</sup>	PATagRFP <sup>15</sup> , Dendra2 <sup>30</sup> and mEos2 <sup>16</sup>	<i>FRET-based biosensors</i>	Venus-mCherry <sup>31</sup>	EYFP-rsTagRFP <sup>19</sup>

**Supplementary Table 3.** Properties of the major chromophore structures observed in fluorescent proteins.

Chemical structure of chromophore	Chromophore name	Number of double bonds in conjugation	Typical absorbance/emission max, nm	Examples of proteins sharing this chromophore
<i>1</i>	non-cyclized tripeptide X65-Tyr66-Gly67	none	none	majority of FPs
<i>3</i>	aromatic $\alpha$ -enolate	2	350-360/ undetectable	PATagRFP (dark state)
<i>4</i>	<b>TagBFP-like</b>	4	400-420/ 456-466	TagBFP, FTs (blue state), PAmCherry (dark state), mCherry-Blue102
<i>5a, 6a, 8</i>	<b>DsRed-like</b> <b>asFP595-like</b> ( <i>anionic</i> )	7, 8	555-610/ 583-650	DsRed, mCherry, PAmCherry ( <i>trans</i> ), TagRFP ( <i>trans</i> ), mKate, PATagRFP, asFP595
<i>5b, 6b</i>	<b>DsRed-like</b> ( <i>neutral</i> )	8	440-463/535, 605-624	mKeima, LSS-mKate1, LSS-mKate2, GmKate
<i>9, 10</i>	<b>mOrange-like, zFP538-like, mKO-like</b> ( <i>anionic</i> )	7	540-548/ 561-565	mOrange, mKO, E2-Orange, zFP538
<i>7</i>	<b>DsRed-like</b> <b>a) H-bonding</b> <b>b) hydrophobic packing</b> <b>c) <math>\pi</math>-<math>\pi</math> stacking</b>	8	>610/ >650	TagRFP657, E2-Crimson, mNeptune, mRouge
<i>11</i>	<b>PSmOrange-like</b>	8	>610/ >650	PSmOrange
<i>12</i>	<b>GFP-like</b> ( <i>anionic</i> )	6	484-493/ 502-510	EGFP, TagGFP2, Emerald
<i>13</i>	<b>GFP-like</b> ( <i>neutral</i> )	6	399/ 511	wtGFP, Sapphire, Ametrine
<i>14</i>	<b>Kaede-like</b>	9	550-580/ 580-595	Kaede, mEos2, mKikGR, Dendra, mIrisFP

See Figure 2 for chemical structures of the chromophores. Additional characteristics of the listed proteins are summarized in the review papers<sup>32-34</sup>.



**Supplementary Table 4.** Amino acid residues in the supporting positions of red fluorescent proteins.

Phenotype/property (structure/transition)		Supporting positions										Ref.		
TagBFP-like (4)		65 H,L	69 K,R	165 A,I,N	179 A	181 A,I							13,35	
Fluorescent timer-like (4→5a,6a)		18 V,L,M		106 A	152 M,I	179 V,A	203 S,I	205 M,L	224 S,A				13,36	
Photoactivatable-like (4→5a,6a)		18 L,M	58 T,A	84 F,W	148 L,S,N,H,T,F	165 V,L,S,A,G,Q	167 P,A,M,G	169 V,T,L	179 T,A,C	207 I,R	218 W,V		14,15,37,38	
Photoswitchable-like (5a↔6a)		146 V,E	150 E	181 F,L,S	183 T								19,20,39,40	
Far-red-shifted- like	H-bonding to N-acylimine oxygen (7a)	63 C,F	64 I,F	165 C									9,10,21,41-45	
	Hydrophobic packing (7c)	181 F,L,V		201 T,V		224 C,N,A								
	π-stacking (7b)	181 F,L,V		201 A,T	224 C,N,A									
Large Stokes shift-like (5b,6b)		69 R,K,Y		148 G,S,N	165 G,A,V,N		167 S,M,A,G,L,N		183 T,S				11,12,46-48	
High quantum yield		145 W	146 E	147 P,A	203 H,I,R								3,5,8,41,49-52	
High photostability		145 W	146 E	147 P,A	163 K								4,8,52	
High pH-stability		145 W	146 E	147 P,A									52-54	
Fast maturation		18 V,L,M	21 T,S	44 M,A	48 A,V	106 I,V, L	152 M,I,L	165 A,I,T	167 M,Q	181 L,V,F	203 S,I,R,H	205 M,L	238 R	8,13,45,55-57
Monomeric state		128 V,T	151 K,T,R, M	153 K,Y	157 A,E		178 I,D,S	180 N,E,S	184 T	198 V,A,I	200 Y,F,N,A		229 S,C,A	6,10,15,56-58

The supporting residues correspond to those on Figure S1 highlighted with yellow color. Residues' numbering follows that for GFP.

## Supplementary Note

### Major steps of molecular evolution

A typical process of the development of fluorescent proteins with required spectral and biochemical properties includes a rational design and several steps of directed molecular evolution (**Fig. 1**). The starting points are design of new spectral phenotype of fluorescent protein based on the known mechanisms of chromophore transformations and choosing of an appropriate template gene.

Next step involves the rational design based on the knowledge of relationship between immediate chromophore environment and postranslational chemistry of fluorescent proteins and photochemical properties of the chromophore. At this step, the selected templates are subjected to multiple site-specific amino acid substitutions at the key and supporting positions, in order to generate a library of site-specific mutants. After screening this library typically in a low-throughput format, several primary clones, possessing the required spectral phenotype, are selected.

Several rounds of directed molecular evolution are then applied to improve the properties of imperfect primary variants. Major challenges to get a successful directed evolution include selecting an optimal biological system for expression, obtaining a sufficient number of clones in a library, and achieving reliable screening conditions. Each round of the directed molecular evolution begins with a generation of a large mutant library, using various strategies of mutagenesis and recombination<sup>59-62</sup>. A high-throughput screening (HTS) of the library and subsequent evaluation of the selected clones in a low-throughput format then follows. The clones exhibiting prominent characteristics of the required phenotype are then applied to the next round of molecular evolution. After several rounds, the screened properties usually reach a plateau, suggesting that either a natural limit is achieved or some bias, caused by the mutagenesis strategy used, is the culprit. To avoid the latter case, a saturated multi-site mutagenesis is typically applied to the selected mutants at amino acid positions found during the rounds of molecular evolution. Finally, testing of advanced variants with optimized characteristics by creating fusion proteins and their imaging in live cells comes before wide usage in the life sciences.

### Historical aspects of the molecular evolution strategy in the development of fluorescent proteins

In the past, fluorescent proteins of new colors such as mFruits have been mainly developed using random mutagenesis on the basis of wild-type protein templates of required weak phenotype<sup>3</sup>. Over time the more structure-to-function information has become available for fluorescent proteins and it resulted into appearance of rational component in the creation of fluorescent proteins. The PAmCherrys<sup>14</sup>, PATagRFP<sup>15</sup>, rsCherrys<sup>20</sup>, rsTagRFP<sup>19</sup>, KikGR<sup>25</sup>, mRouge<sup>41</sup> and series of TagBFP-<sup>35</sup> and LSS-like<sup>46</sup> fluorescent proteins have different colors and photochemical behavior and all of them were developed using rational mutagenesis at amino acids around chromophore to find weak phenotype followed by random mutagenesis for the further improvement of mainly brightness, maturation and photostability.

### Correlation between residues at key positions and fluorescent protein phenotype

The residues at key positions 84, 148 and 203 for TagBFP-like phenotype mainly sterically hinder Tyr-ring of the chromophore occupies position necessary for the oxidation of

C $\alpha$ 2-C $\beta$ 2 double bond between Tyr-ring and imidazolinone and improve brightness of blue chromophore 4.

The residues at key positions 69 and 84 for FT-like phenotype slow down 4-to-5a,6a transition and increase brightness of blue chromophore 4. The residues at crucial positions 69 and 203 for photoactivatable-like phenotype block 4-to-5a,6a transition. The residues at main positions 148, 165, 167 and 203 for photoswitchable-like phenotype provide space around Tyr-ring of the chromophore and ensure flexibility of the chromophore for *cis-trans* isomerization of the chromophore to occur.

The far-red shift in RFPs can be attributed by the formation of hydrogen bond between residues at positions 14, 16 and 44 and N-acylimine oxygen, by hydrophobic parking around chromophore provided by hydrophobic residues at positions 84, 148, 165 and 167 and by stacking interactions between aromatic residue at positions 65, 69, 148 and 203 and Tyr-ring of the chromophore.

Large Stokes shift-like phenotype relies on the Glu and Asp acidic amino acids at positions 148, 165 and 167 which stabilize neutral chromophore 5b,6b and support proton transfer in excited state.

High photostability of RFPs mainly depends on amino acids at positions 64, 99 and 165 which provide tight packing around chromophore and block photochemical reactions.

Fast maturation of RFPs is primarily affected by residues at positions 42, 69, 179, and 224 which facilitate 2-to-5a,6a transition.

The A-B and A-C dimerizing interfaces in RFPs are mainly disrupted by charged amino acid residues at positions 126, 162, 166 and 168.

High quantum yield in RFPs depends on the residues at positions 70, 148 and 167 which provide planar configuration and rigidity of the chromophore.

High pH-stability of blue and red chromophores 4 and 5a,6a, respectively, is provided by the residues at positions 96 and 167 which stabilize negative charge on CO-group of imidazolinone and Tyr-ring, respectively.

Maturation rate, pH-stability, quantum yield and photostability may be affected by the width of the water-filled pore which regulated by the residues at positions 145-147.

There are many examples in the world of fluorescent proteins where a particular feature (e.g., a red color in DsRed) is a synergistic effect of a large number of residue substitutions, including many residues that are remote from the active site. How these positions control fluorescence via coupling to each other is poorly understood. This is why the directed evolution approach will be useful for final turning of the residues remote from chromophore with found spectral phenotype.

## Supplementary references

1. Karasawa, S., Araki, T., Nagai, T., Mizuno, H. & Miyawaki, A. Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer. *Biochem. J.* **381**, 307-312 (2004).
2. Sakaue-Sawano, A. et al. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487-498 (2008).
3. Shaner, N.C. et al. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567-1572 (2004).
4. Shaner, N.C. et al. Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5**, 545-551 (2008).
5. Merzlyak, E.M. et al. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. *Nat. Methods* **4**, 555-557 (2007).
6. Kredel, S. et al. mRuby, a bright monomeric red fluorescent protein for labeling of subcellular structures. *PLoS One* **4**, e4391 (2009).
7. Wang, L., Jackson, W.C., Steinbach, P.A. & Tsien, R.Y. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc. Natl. Acad. Sci. USA* **101**, 16745-16749 (2004).
8. Shcherbo, D. et al. Far-red fluorescent tags for protein imaging in living tissues. *Biochem. J.* **418**, 567-574 (2009).
9. Shcherbo, D. et al. Near-infrared fluorescent proteins. *Nat. Methods* **7**, 827-829 (2010).
10. Lin, M.Z. et al. Autofluorescent proteins with excitation in the optical window for intravital imaging in mammals. *Chem. Biol.* **16**, 1169-1179 (2009).
11. Kogure, T. et al. A fluorescent variant of a protein from the stony coral *Montipora* facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. *Nat. Biotechnol.* **24**, 577-581 (2006).
12. Piatkevich, K.D. et al. Monomeric red fluorescent proteins with a large Stokes shift. *Proc. Natl. Acad. Sci. USA* **107**, 5369-5374 (2010).
13. Subach, F.V. et al. Monomeric fluorescent timers that change color from blue to red report on cellular trafficking. *Nat. Chem. Biol.* **5**, 118-126 (2009).
14. Subach, F.V. et al. Photoactivatable mCherry for high-resolution two-color fluorescence microscopy. *Nat. Methods* **6**, 153-159 (2009).
15. Subach, F.V., Patterson, G.H., Renz, M., Lippincott-Schwartz, J. & Verkhusha, V.V. Bright monomeric photoactivatable red fluorescent protein for two-color super-resolution sptPALM of live cells. *J. Am. Chem. Soc.* **132**, 6481-6491 (2010).
16. McKinney, S.A., Murphy, C.S., Hazelwood, K.L., Davidson, M.W. & Looger, L.L. A bright and photostable photoconvertible fluorescent protein. *Nat. Methods* **6**, 131-133 (2009).
17. Habuchi, S., Tsutsui, H., Kochaniak, A.B., Miyawaki, A. & van Oijen, A.M. mKikGR, a monomeric photoswitchable fluorescent protein. *PLoS One* **3**, e3944 (2008).
18. Chudakov, D.M. et al. Kindling fluorescent proteins for precise in vivo photolabeling. *Nat. Biotechnol.* **21**, 191-194 (2003).
19. Subach, F.V. et al. Red fluorescent protein with reversibly photoswitchable absorbance for photochromic FRET. *Chem. Biol.* **17**, 745-755 (2010).

20. Stiel, A.C. et al. Generation of monomeric reversibly switchable red fluorescent proteins for far-field fluorescence nanoscopy. *Biophys. J.* **95**, 2989-2997 (2008).
21. Morozova, K.S. et al. Far-red fluorescent protein excitable with red lasers for flow cytometry and superresolution STED nanoscopy. *Biophys. J.* **99**, L13-15 (2010).
22. Tsutsui, H., Karasawa, S., Okamura, Y. & Miyawaki, A. Improving membrane voltage measurements using FRET with new fluorescent proteins. *Nat. Methods* **5**, 683-685 (2008).
23. Drobizhev, M., Tillo, S., Makarov, N.S., Hughes, T.E. & Rebane, A. Absolute two-photon absorption spectra and two-photon brightness of orange and red fluorescent proteins. *J. Phys. Chem. B* **113**, 855-859 (2009).
24. Nienhaus, K., Nienhaus, G.U., Wiedenmann, J. & Nar, H. Structural basis for photo-induced protein cleavage and green-to-red conversion of fluorescent protein EosFP. *Proc Natl Acad Sci USA* **102**, 9156-9159 (2005).
25. Tsutsui, H., Karasawa, S., Shimizu, H., Nukina, N. & Miyawaki, A. Semi-rational engineering of a coral fluorescent protein into an efficient highlighter. *EMBO Rep.* **6**, 233-238 (2005).
26. Stepanenko, O.V. et al. Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, "dimer2", and DsRed1. *Biochemistry* **43**, 14913-14923 (2004).
27. Kedrin, D. et al. Intravital imaging of metastatic behavior through a mammary imaging window. *Nat. Methods* **5**, 1019-10121 (2008).
28. Strack, R.L. et al. A noncytotoxic DsRed variant for whole-cell labeling. *Nat. Methods* **5**, 955-957 (2008).
29. Strack, R.L. et al. A Rapidly Maturing Far-Red Derivative of DsRed-Expres2 for Whole-Cell Labeling. *Biochemistry* **48**, 8279-8281 (2009).
30. Baker, S.M., Buckheit, R.W., 3rd & Falk, M.M. Green-to-red photoconvertible fluorescent proteins: tracking cell and protein dynamics on standard wide-field mercury arc-based microscopes. *BMC Cell Biol.* **11**, 15 (2010).
31. Hazelwood, K.L. et al. Searching the fluorescent protein color palette for new FRET pairs. *Proc. SPIE* **6868**, 68681-686812 (2008).
32. Davidson, M.W. & Campbell, R.E. Engineered fluorescent proteins: innovations and applications. *Nat. Methods* **6**, 713-717 (2009).
33. Chudakov, D.M., Matz, M.V., Lukyanov, S. & Lukyanov, K.A. Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol. Rev.* **90**, 1103-1163 (2010).
34. Piatkevich, K.D., Efremenko, E.N., Verkhusha, V.V. & Varfolomeev, S.D. Red fluorescent proteins and their properties. *Russ. Chem. Rev.* **79**, 243-258 (2010).
35. Subach, O.M. et al. Conversion of red fluorescent protein into a bright blue probe. *Chem. Biol.* **15**, 1116-1124 (2008).
36. Terskikh, A. et al. "Fluorescent timer": protein that changes color with time. *Science* **290**, 1585-1588 (2000).
37. Subach, F.V. et al. Photoactivation mechanism of PAmCherry based on crystal structures of the protein in the dark and fluorescent states. *Proc. Natl. Acad. Sci. USA* **106**, 21097-21102 (2009).
38. Gunewardene, M.S. et al. Superresolution Imaging of Multiple Fluorescent Proteins with Highly Overlapping Emission Spectra in Living Cells. *Biophys. J.* **101**, doi:10.1016/j.bpj.2011.1007.1049 (2011).

39. Chudakov, D.M., Feofanov, A.V., Mudrik, N.N., Lukyanov, S. & Lukyanov, K.A. Chromophore environment provides clue to "kindling fluorescent protein" riddle. *J. Biol. Chem.* **278**, 7215-7219 (2003).
40. Wiedenmann, J. et al. From EosFP to mIrisFP: structure-based development of advanced photoactivatable marker proteins of the GFP-family. *J. Biophotonics* **4**, 377-390 (2011).
41. Chica, R.A., Moore, M.M., Allen, B.D. & Mayo, S.L. Generation of longer emission wavelength red fluorescent proteins using computationally designed libraries. *Proc. Natl. Acad. Sci. USA* **107**, 20257-20262 (2010).
42. Shu, X., Shaner, N.C., Yarbrough, C.A., Tsien, R.Y. & Remington, S.J. Novel chromophores and buried charges control color in mFruits. *Biochemistry* **45**, 9639-9647 (2006).
43. Shu, X., Wang, L., Colip, L., Kallio, K. & Remington, S.J. Unique interactions between the chromophore and glutamate 16 lead to far-red emission in a red fluorescent protein. *Protein Sci.* **18**, 460-466 (2009).
44. Strack, R.L. et al. A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* **48**, 8279-8281 (2009).
45. Kredel, S. et al. Optimized and far-red-emitting variants of fluorescent protein eqFP611. *Chem. Biol.* **15**, 224-233 (2008).
46. Piatkevich, K.D., Malashkevich, V.N., Almo, S.C. & Verkhusha, V.V. Engineering ESPT pathways based on structural analysis of LSSmKate red fluorescent proteins with large Stokes shift. *J. Am. Chem. Soc.* **132**, 10762-10770 (2010).
47. Henderson, J.N. et al. Excited state proton transfer in the red fluorescent protein mKeima. *J. Am. Chem. Soc.* **131**, 13212-13213 (2009).
48. Kogure, T., Kawano, H., Abe, Y. & Miyawaki, A. Fluorescence imaging using a fluorescent protein with a large Stokes shift. *Methods* **45**, 223-226 (2008).
49. Lukyanov, K.A. et al. Natural animal coloration can be determined by a nonfluorescent green fluorescent protein homolog. *J. Biol. Chem.* **275**, 25879-25882 (2000).
50. Pletnev, S., Subach, F.V., Dauter, Z., Wlodawer, A. & Verkhusha, V.V. Structural basis for reversible photoswitching of absorbance spectra in red fluorescent protein rsTagRFP. *Submitted*.
51. Gurskaya, N.G. et al. GFP-like chromoproteins as a source of far-red fluorescent proteins. *FEBS Lett.* **507**, 16-20 (2001).
52. Stiel, A.C. et al. 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. *Biochem. J.* **402**, 35-42 (2007).
53. Wilmann, P.G. et al. The 2.1 Å crystal structure of copGFP, a representative member of the copepod clade within the green fluorescent protein superfamily. *J. Mol. Biol.* **359**, 890-900 (2006).
54. Bravaya, K.B., Korovina, N., Subach, O.M., Verkhusha, V.V. & Krylov, A.I. Puzzling over the mechanism of red chromophore maturation in the DsRed-like proteins: The nature of the blue intermediate revealed. *Submitted*.
55. Pletnev, S., Subach, F.V., Dauter, Z., Wlodawer, A. & Verkhusha, V.V. Understanding blue-to-red conversion in monomeric fluorescent timers and hydrolytic degradation of their chromophores. *J. Am. Chem. Soc.* **132**, 2243-2253 (2010).
56. Campbell, R.E. et al. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA* **99**, 7877-7882 (2002).

57. Wiedenmann, J. et al. A far-red fluorescent protein with fast maturation and reduced oligomerization tendency from *Entacmaea quadricolor* (Anthozoa, Actinaria). *Proc. Natl. Acad. Sci. USA* **99**, 11646-11651 (2002).
58. Nienhaus, K., Nar, H., Heilker, R., Wiedenmann, J. & Nienhaus, G.U. Trans-cis isomerization is responsible for the red-shifted fluorescence in variants of the red fluorescent protein eqFP611. *J. Am. Chem. Soc.* **130**, 12578-12579 (2008).
59. Wang, T.W. et al. Mutant library construction in directed molecular evolution: casting a wider net. *Mol. Biotechnol.* **34**, 55-68 (2006).
60. Farinas, E.T., Bulter, T. & Arnold, F.H. Directed enzyme evolution. *Curr. Opin. Biotechnol.* **12**, 545-551 (2001).
61. Yuan, L., Kurek, I., English, J. & Keenan, R. Laboratory-directed protein evolution. *Microbiol. Mol. Biol. Rev.* **69**, 373-392 (2005).
62. Giver, L. & Arnold, F.H. Combinatorial protein design by in vitro recombination. *Curr. Opin. Chem. Biol.* **2**, 335-338 (1998).