

## **Supplementary Material**

### **Mammalian-Specific Sequences in *Pou3f2* Contribute to Maternal Behaviour**

Makoto Nasu, Saori Yada, Atsushi Igarashi, Den'etsu Sutoo, Kayo Akiyama, Meguru Ito, Nobuaki Yoshida, and Shintaroh Ueda

Supplementary table and figure legends

Table S1

Sequence information of vertebrate species used in supplementary fig. S1.

Species	Latin name	Transcript ID
Human	<i>Homo sapiens</i>	NM 005604.2
Chimpanzee	<i>Pan troglodytes</i>	AB554076
Gorilla	<i>Gorilla gorilla</i>	AB554075
Orangutan	<i>Pongo pygmaeus</i>	AB554074
Cow	<i>Bos taurus</i>	NM 001080348.1/ENSBTAT00000034488
Lesser hedgehog tenrec	<i>Echinops telfairi</i>	ENSETET00000020341
Mouse	<i>Mus musculus</i>	NM 008899.1/ENSMUST00000064587
Rat	<i>Rattus norvegicus</i>	NM 172085.1/ENSRNOT00000009043
Opossum	<i>Monodelphis domestica</i>	ENSMODT00000022991
Japanese quail	<i>Coturnix japonica</i>	AF091043.1
Green anole	<i>Anolis carolinensis</i>	AB001869.1
<i>Xenopus laevis</i>	<i>Xenopus laevis</i>	NM 001101755.1
<i>Xenopus tropicalis</i>	<i>Xenopus tropicalis</i>	AB557604
Medaka	<i>Oryzias latipes</i>	ENSORLT00000002764
Stickleback	<i>Gasterosteus aculeatus</i>	ENSGACT00000006578
Takifugu	<i>Takifugu rubripes</i>	ENSTRUT00000038615
Tetraodon	<i>Tetraodon nigroviridis</i>	ENSTNIT00000016300
Zebrafish	<i>Danio rerio</i>	NM 131160.1/ENSDART00000113914

Figure S1.

Multiple alignment of POU3F2s among *Vertebrata*.

For multiple alignment, we used MAFFT on a CLC Bio Sequence Viewer. Sequence data contained nine mammals, one bird, one reptile, two amphibians, and five fish (See supplementary table S1). Red boxes indicate homopolymeric amino acid repeats, polyG, polyQ, and polyP. Blue boxes indicate DNA binding domain, POU-specific domain, and POU-homeodomain. Sequence gaps are indicated by the hyphen.

Figure S2.

Comparison between  $+/+$  and *tro/tro* mice.

(A) Immunohistochemical distribution of POU3F2.

(B) Body weight.

(C) Brain weight per body weight.

(D) Nissl staining of brain coronal section around hypothalamus.

(E) Nissl staining of brain sagittal section.

Abbreviations: Cx, cerebral cortex; PVN, paraventricular nucleus; SO, supraoptic nucleus; DG, dentate gyrus; CA1, cornet d'Ammon; ST, neostriatum; DR, dorsal raphe nucleus; SCh, suprachiasmatic nucleus; 3V, third ventricle; LV, lateral ventricle; 4V, fourth ventricle.

Figure S3.

Comparison among  $+/+$ ,  $+/Δ$ , and  $Δ/Δ$  mice.

(A) Relative expression level of *Pou3f2* by qPCR.

Each data is indicated as a ratio to the expression level of  $+/+$  mice. No statistical differences were observed among  $+/+$ ,  $+/Δ$ , and  $Δ/Δ$  mice.

(B) Genotype ratio of pups from  $+/Δ$  parents.

(C) Pups viability. Black, meshed red, and red bar represent  $+/+$ ,  $+/Δ$ , and  $Δ/Δ$ , respectively.

\*\* $P < 0.01$ .

Figure S4.

Confocal images of tryptophan hydroxylase 2 immunohistochemistry in the raphe nucleus.

(A)  $+/+$  (left) and *tro/tro* mice (right) in the raphe nucleus.

(B)  $+/+$  (left) and  $Δ/Δ$  mice (right) in the raphe nucleus.

(C) Atlas of the raphe nucleus.

Abbreviations: Aq, aqueduct; DRD, dorsal part of the dorsal raphe nucleus; DRV, ventral

part of the dorsal raphe nucleus; DRL, lateral part of the dorsal raphe nucleus.

Figure S5.

Predicted profile of the intrinsically disordered structure of mouse POU3F2, mouse POU3F2  $\Delta$ GQP, and *Xenopus tropicalis* POU3F2.

This analysis was performed on the PrDOS server to predict intrinsically disordered regions of the protein from the amino acid sequence (<http://prdos.hgc.jp/cgi-bin/top.cgi/>) (Ishida and Kinoshita 2007). Many proteins have regions with very flexible and unstable structures, referred to as intrinsically disordered, even in their native states. Disordered regions are involved in many biological processes, such as regulation, signalling, and cell cycle control. The primary role of intrinsically disordered regions seems to be the molecular recognition of proteins or DNA. Upon binding with ligands, disorder-to-order transitions are frequently observed, where the flexibility of disordered regions may be necessary to interact with multiple partners with high-specificity and low-affinity. Therefore, the prediction of disordered regions in proteins is important for the functional annotation of proteins. PrDOS: prediction of disordered protein regions from amino acid sequence.

(A-C) Transition of sequential disordered probability of amino acid sequences; mouse POU3F2 (A), POU3F2  $\Delta$ GQP form (B), and *X. tropicalis* POU3F2 (C). Prediction false positive rate was set to 5.0%. Orange boxes indicate homopolymeric amino acid repeats and orange lines indicate the positions where homopolymeric amino acid repeats are inserted in mammals. Blue boxes represent POU-specific domain and POU-homeodomain.

(D-F) Properties of disordered probability in sequences. Amino acids with high disordered probability (>0.5 threshold) are highlighted in red. Yellow boxes indicate homopolymeric amino acid repeats and yellow lines indicate the positions where homopolymeric amino acid repeats are inserted in mammals. Blue boxes represent POU-specific domain and POU-homeodomain.

In POU3F2, most of the homopolymeric amino acid repeats showed high disordered probability. Therefore, disordered values of the sequence without homopolymeric amino acid repeats were comparatively low in the respective regions in POU3F2  $\Delta$ GQP and xPOU3F2. Moreover, the disordered probability of N-terminal 65 residues of xPOU3F2 is low in comparison with mouse POU3F2 and POU3F2  $\Delta$ GQP, in which homopolymeric glycine repeats are uniquely present in mammals. Therefore, this difference between POU3F2  $\Delta$ GQP and xPOU3F2 might cause prominent phenotype differences in behavioural tests and monoamine neurotransmission between *tro/tro* and  $\Delta/\Delta$  mice. In mammalian



POU3F2, the acquisition of not only homopolymeric amino acid repeats but also amino acid substitutions in transactivation domains might contribute to mammal-specific function and expression.

### **Reference**

Ishida T, Kinoshita K. 2007. PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res.* 35: W460-W464.

20 40 60 80 100 120

Mammalia  
 Human MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Chimpanzee MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Gorilla MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Orangutan MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Mouse MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Rat MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Cow MATAA5NHYLLT--SSASIVHTEPGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Tenrec MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ110  
 Opossum MATAA5NHYLLT--SSASIVHAEPGGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SPWSTS--PLGQPDIKPSVVVQ111  
 Quail MATAA5NHYLLA--SGSPMVAHEPGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--APWAAAAAAGALGPPDIKPA--AVQ112  
 Green-anole MATTA5NHYLLA--SGSPMVAHEPGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--SGDSSPWSTS--PLGQPDIKPSVVQ107  
 X.tropicalis MATTA5NHYLLA--SGSPMVAHEPGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGGGGCGGGGGGGGGGGGGD--APWAAAAAAGALGPPDIKPA--AVQ112  
 Amphibia  
 X.tropicalis MATTA5NHYLLG--SGSIVHAD--PGCMQOAGGYYREASLVQDGYALQSNHPLSHAHQWITALSHGGD--APWAST--PLGQDDIQT--VQ85  
 Fishes  
 Zebrafish MATTA5NHYLLT--SPSIVHSE--PGSMQOAT--AYRDAQTLQDYSLQSNHPLSHAHQWITALSHGGEG--GPWSS--PLGQDDIKPA--VQ85  
 Takifugu MATAA5NHYVLTTPSAPPHSE--SGSMQOAA--AYRDAQTLQDYSLTPGGHPLSHAHQWITALSHGDAG--APWSSP--PLCEQDQKPA--VL88  
 Tetraodon MATAA5NHYVLTTPSAPPHSE--SGSMQOAA--AYRDAQTLQDYSLTPGGHPLSHAHQWITALSHGDGT--APWSSP--PLGQDDQKPA--VL88  
 Medaka MATAA5NHYSLLT--SSASIVHSE--ACSMQOAA--AYRDAQTLQDYSLQSNHPLSHAHQWITALSHGEG--APWSSP--PLCEAQDQKPA--VQ86  
 Stickleback MATAA5NHYLLT--SSASIVHSE--PGSMQOAT--AYRDAQTLQDYSLQSNHPLSHAHQWITALSHGEG--VPWSSS--PLCEAQDQKPA--VQ86

140 160 180 200 220 240

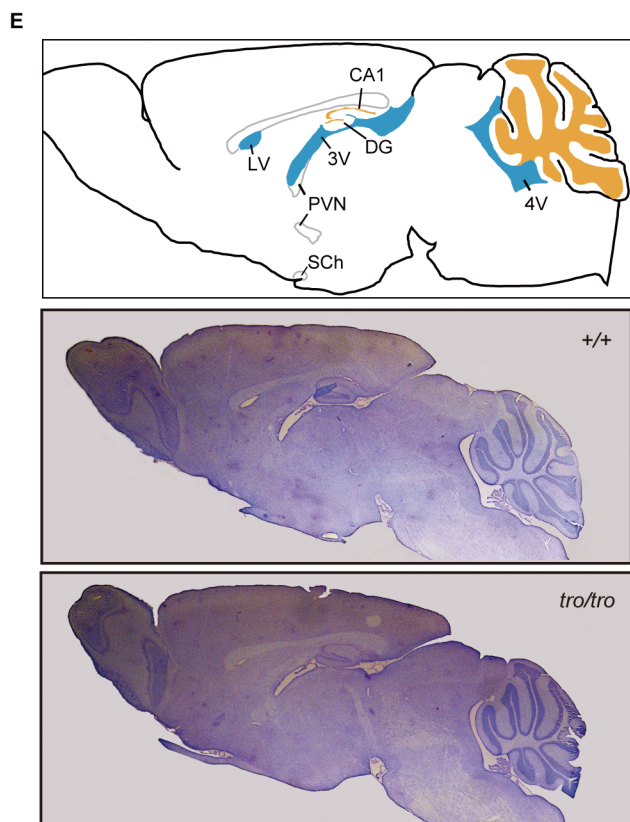
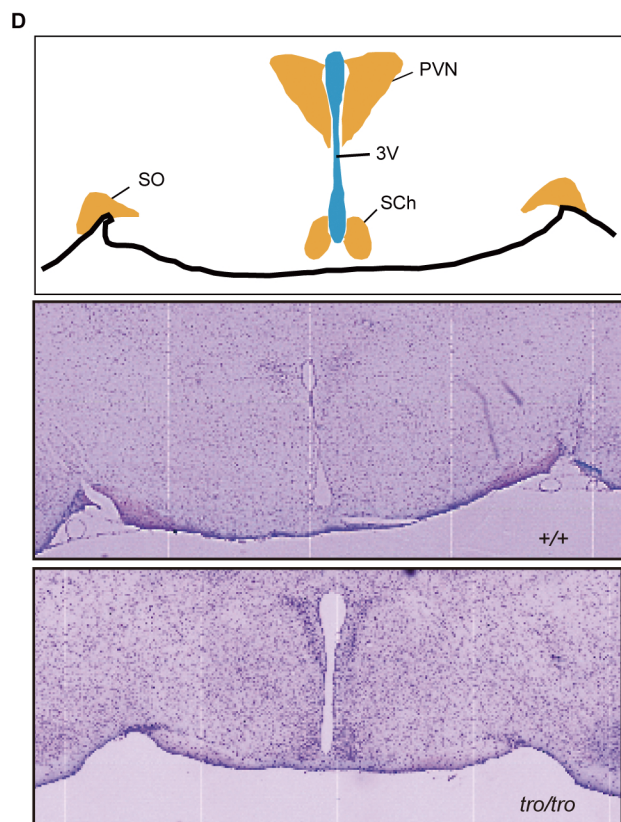
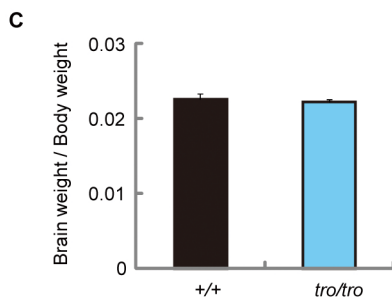
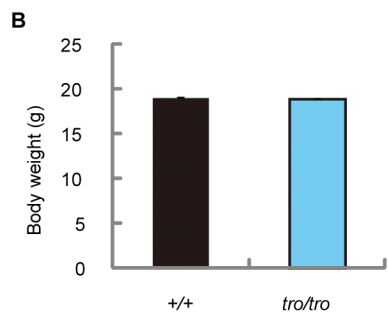
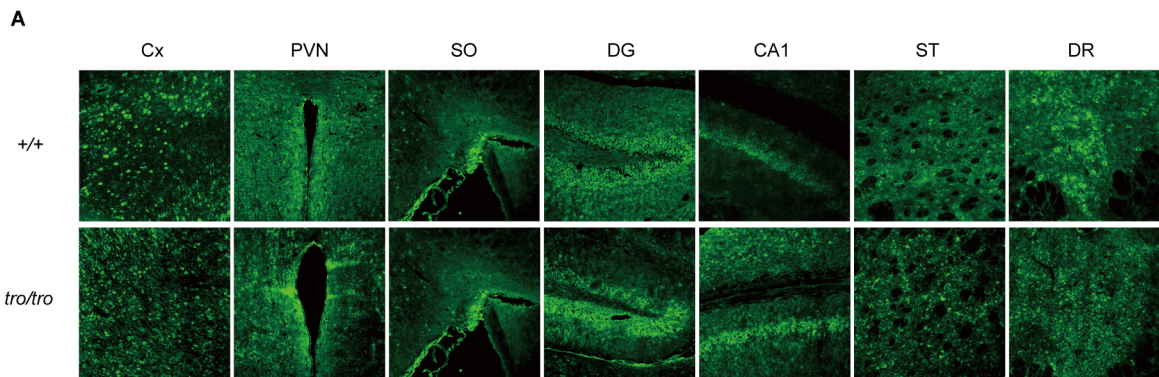
Mammalia  
 Human QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH219  
 Chimpanzee QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH219  
 Gorilla QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH219  
 Orangutan QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH219  
 Mouse QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH221  
 Rat QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH219  
 Cow QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH221  
 Tenrec QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH217  
 Opossum QCGRDELHGPGALQQO--QQOQQOQQOQQOQQOQQOQQRPHLVH--AANHPPGCG--AWRSAAAAAHLPPSMGASN--GGLLYSQ--PSFTVNGML--GAGGQPAGLHHGLRDAH219  
 Quail AAPRGDLPPP--PQHPPPPGAPHLVHGGGGGHH--AWRA--GAAHLLPPGMAAANGAQQALLYPO--PPGFTVNGML--GA--AQP--ALHHHGLRDAH203  
 Green-anole ACGRDDL--QQOQQOQQOQQOQQOQQOQQRPHLVH--ACGSHAAAAVAAVART--GSAHLPPGMAAANGAQQALLYSQ--PPGFTVNGML--GS--QP--GMHHGLRDAH203  
 X.tropicalis S--RDELHGA--G--TLOHQGRAPHLVH--AHGNHGGP--AWRS--TSGAHL--SMASSNG--GGLLYSQ--PSFTVNGMIPGCS--GQ--GMHHGLRDAH169  
 Fishes  
 X.laieus S--RDELHVS--G--TLOHQGRAPHLVH--AHGNHGGP--AWRS--TSGTSL--SMASSNG--GGLLYSQ--PSFTVNGMIPGCS--GQ--GIHHGLRDS169  
 Zebrafish HDSDRELLONS--SSLS--QQOQQOQQOQQOQQRPHVPAHQ--QAHQDAR--AWRTSTASTH--PGMTTSE--QSLVSYQ--SGF--GLM--PG--GEOGGMHHLRDEE181  
 Takifugu HDSDRELLONS--SSLS--QQOQQOQQOQQOQQRPHVPAHQ--QAHQDAR--AWRTSTASTH--PGMTTSE--QSLVSYQ--SGF--GLM--PG--GEOGGMHHLRDEE181  
 Tetraodon HDSDRELLONS--SSLQPPQQOQQOQQOQQRPHVPAHQ--QPHQDAR--AWRTSTASTH--PGMTTSE--QSLVSYQ--SAF--GLM--PG--GEOGGMHHLRDEE185  
 Medaka QT--RDEMHS--SNLQHQQRPHLVH--THGNHGDGR--AWRT--TAAHI--PSMTTNG--QSLIYSQ--PGFNVGLI--PGCSGQ--GMHHHLRDS170  
 Stickleback GA--RDEMHS--SNLQHQQRPHLVH--THGNHGDGR--AWRT--TAAHI--PSMTTNG--QSLIYSQ--PGFNVGLI--PGCSGQ--GMHHHLRDS171

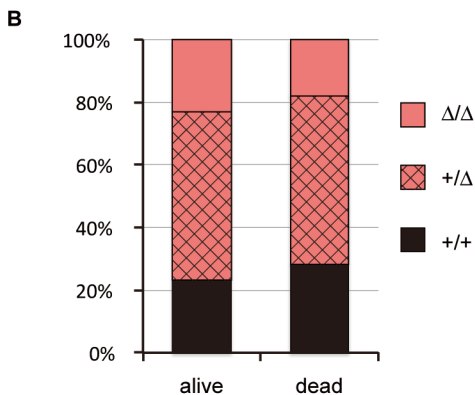
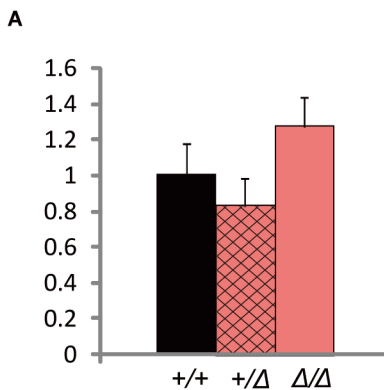
260 280 300 320 340 360

Mammalia  
 Human DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP327  
 Chimpanzee DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP327  
 Gorilla DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP327  
 Orangutan DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP327  
 Mouse DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP327  
 Rat DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP327  
 Cow DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP329  
 Tenrec DE---PHADHHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP328  
 Opossum DE---QHNGEHPHPSHPHQ--PPPPPPGGPPGHC---AHDHPHSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP325  
 Quail DE---APGPPMPHGAHFG--GHPGAPGAGCGGGGAAAGGPH--AHS--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP315  
 Green-anole DE---PPPPPPPPHPSHPHQ--QQOQQOQQOQQOQQOQQOQQRPHLVH--GHH--HSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP326  
 X.tropicalis DD---HH--GE--HGHQPPTQOQHSLQCG--GHHD--HSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP326  
 Amphibia  
 X.laieus DD---HH--GD--HGHQVQSAQQOHSLQCG--GHOD--HSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP326  
 Fishes  
 Zebrafish DD---HHS--PLSD--HGH--PSPHQHQS--SHHD--HSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP326  
 Takifugu HH---HHS--PLSE--HGGGPAHQOPLSHHHQHG--GHHD--QSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP328  
 Tetraodon HH---HHS--PLSE--HGGGPAHQOPLSHHHQHG--GHHD--QSD--EDTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP328  
 Medaka EE---HHS--PLSD--HGH--PASQHQHQR--SHHD--HSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP328  
 Stickleback DD---HHS--PLSD--HGH--PASQHQHQR--SHHD--HSD--EDTPTSDDELEQFAKQFKRIRKLGFTQADVGLALGTLYGNVFSQTTICRFEALQ5FNKMCKLKP329

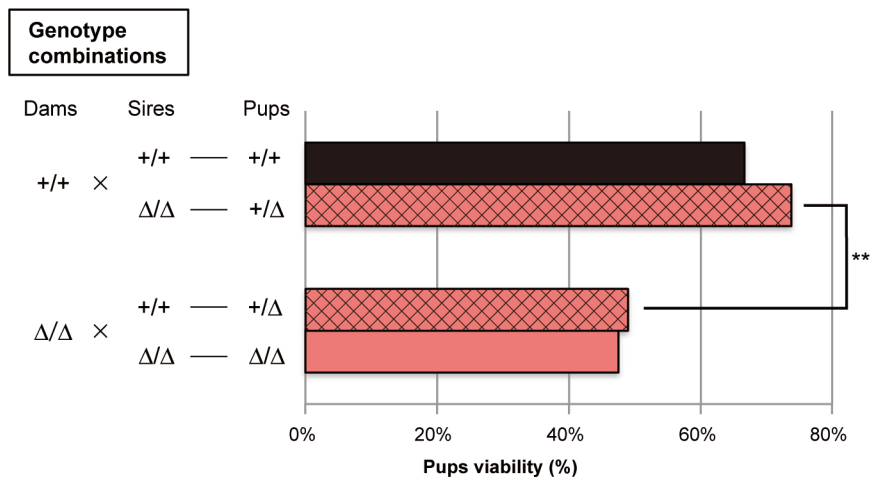
380 400 420 440 460 480

Mammalia  
 Human LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--443  
 Chimpanzee LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--443  
 Gorilla LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--443  
 Orangutan LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--443  
 Mouse LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--445  
 Rat LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--445  
 Cow LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--444  
 Tenrec LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--444  
 Opossum LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--388  
 Quail LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--437  
 Green-anole LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--441  
 Amphibia  
 X.tropicalis LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--SVQ--382  
 Fishes  
 X.laieus LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--SVQ--382  
 Zebrafish LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKSAQEIISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--443  
 Takifugu LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKGAAEISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--396  
 Tetraodon LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKGAAEISLADSLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--400  
 Medaka LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKPAASIEISLADNLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--381  
 Stickleback LKNWLEADSSGSPSTSDIKIAAQRKRKRRTSIEVSKGALSHFLKCKPKPAASIEISLADNLSLEQEVVRRVWFNCNRQREKRTMPGGTLPGAEDVYGSRDT--PHHCVGT--PVQ--382

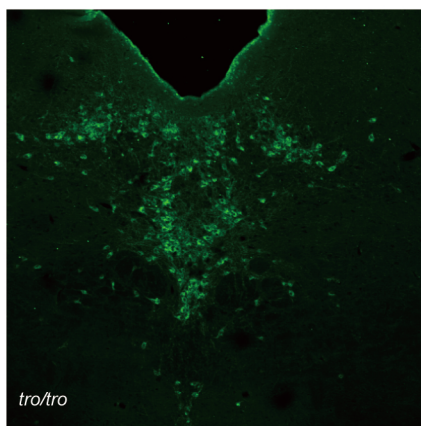
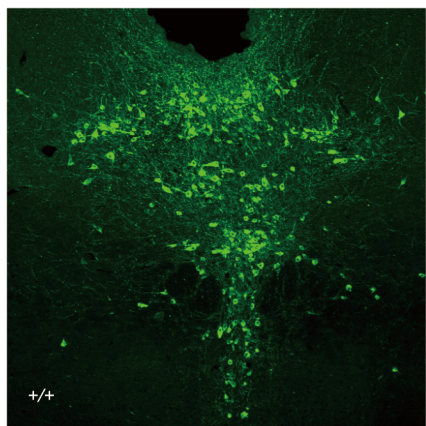
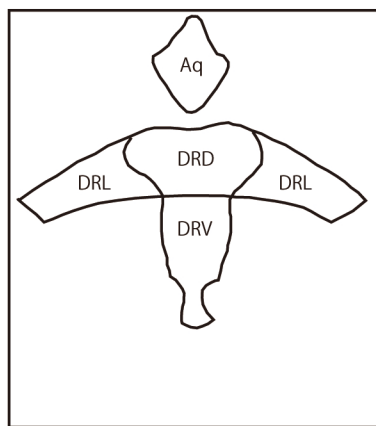
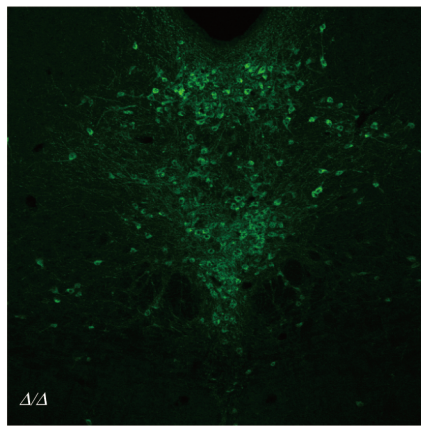
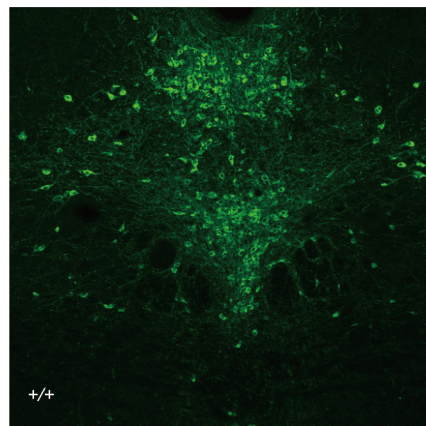




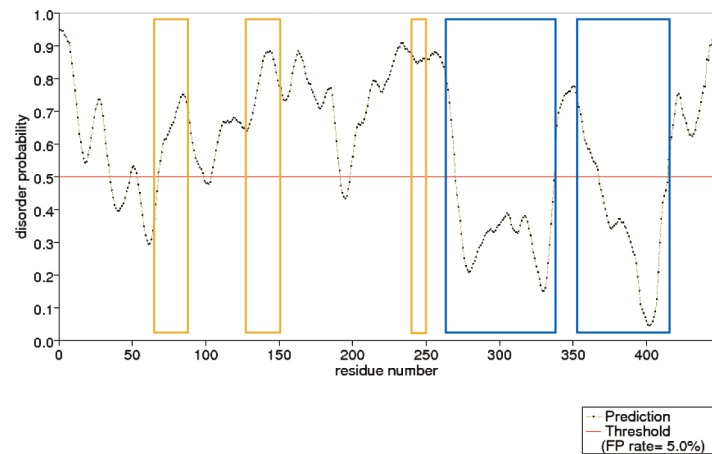
**C**



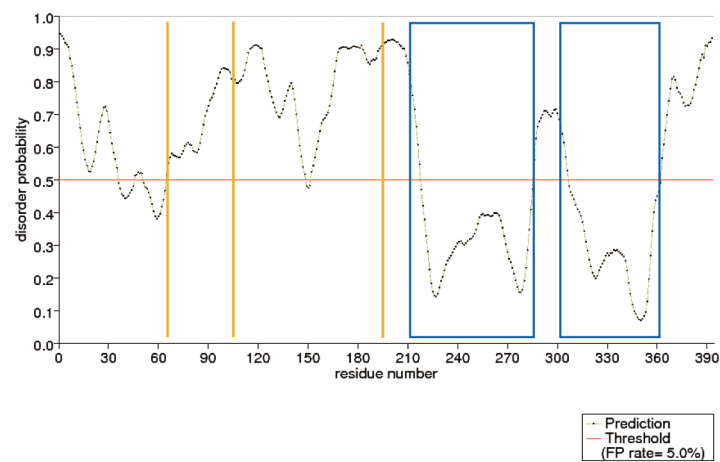


**A****C****B**

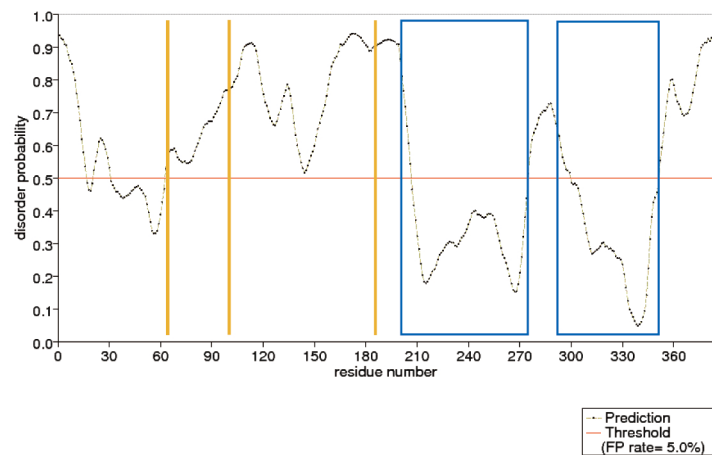
**A** *Mus musculus* POU3F2\_wild-type



**B** *Mus musculus* POU3F2\_ΔGQP form



**C** *Xenopus tropicalis* POU3F2\_wild-type



**D** *Mus musculus* POU3F2\_wild-type

1	<b>MATAASNHYS</b>	<b>LTSSASIVH</b>	<b>AEPFGMQQG</b>	<b>AGGY</b> REAQSL	VQGDY <b>GALQS</b>	50
51	NGHPLSHAHQ	WITALSH <b>GGG</b>	<b>GGGGGGGGG</b>	<b>GGGGGGGGD</b>	<b>SPWSTSP</b> LGGQ	100
101	<b>PDIKPSVVVQ</b>	<b>QGGRGDELHG</b>	<b>PGALQQQHQQ</b>	<b>QQQQQQQQQ</b>	<b>QQQQQQQQQ</b>	150
151	<b>QRPPHLVHHA</b>	<b>ANHHPPGAW</b>	<b>RSAAAAAHP</b>	<b>PSMGASNGGL</b>	<b>LYSQPSFTVN</b>	200
201	<b>GMLGAGQPPA</b>	<b>GLHHHGLRDA</b>	<b>HDEPHHADHH</b>	<b>PHPHSHPHQQ</b>	<b>PPPPPPQGP</b>	250
251	<b>PGHPGAHHDP</b>	<b>HSEDTPTS</b> D	DLEQFAKQFK	QRRIKLGFQ	ADVGLALGTL	300
301	<b>YGNVFSQTTI</b>	CRFEALQLSF	KNMCKLPLL	NKWLEED <b>SS</b>	<b>SGSPTSIDKI</b>	350
351	<b>AAQGRKRKRK</b>	<b>TSIEVSVKGA</b>	LESHFLKCPK	PSAQEITSLA	DSLQLEKEVV	400
401	<b>RVWFCNRRQK</b>	<b>EKRMT</b> PPGGT	<b>LPGAEDVYGG</b>	<b>SRDTPPHGCV</b>	<b>QTPVQ</b>	450

**E** *Mus musculus* POU3F2\_ΔGQP form

1	<b>MATAASNHYS</b>	<b>LTSSASIVH</b>	<b>AEPFGMQQG</b>	<b>AGGY</b> REAQSL	VQGDY <b>GALQS</b>	50
51	NGHPLSHAHQ	WITAL <b>SDGS</b>	<b>PWSTSP</b> LGGQ	<b>DIKPSVVVQ</b>	<b>GGRGDELHGP</b>	100
101	<b>GALQHQRPPH</b>	<b>L VHHAANHH</b>	<b>GPGAWRSAAA</b>	<b>AAHLPPSMGA</b>	<b>SNGGLLYS</b> QP	150
151	<b>SFTVNGMLGA</b>	<b>GGQPAGLHHH</b>	<b>GLRDAHDEPH</b>	<b>HADHHPHPHS</b>	<b>HPHQ</b> Q <b>GGPPG</b>	200
201	<b>HPGAHHDPHS</b>	<b>DEDTPTS</b> DDL	EQFAKQFKQR	RIKLGFTQAD	VGLALGTLYG	250
251	<b>NVFSQTTICR</b>	FEALQLSFKN	MCKLPLLNK	WLEED <b>SSSG</b>	<b>SPTSIDKIAA</b>	300
301	<b>QGRKRKRKRTS</b>	IEVSVKGALE	SHELKCPKPS	AQEITSLADS	LQLEKEVVRV	350
351	<b>WFCNRRQKEK</b>	<b>RMT</b> PPGGTLP	<b>GAEDVYGGSR</b>	<b>DTPPHGVQT</b>	<b>PVQ</b>	400

**F** *Xenopus tropicalis* POU3F2\_wild-type

1	<b>MATTASNHYN</b>	<b>LLSGSS</b> SIVH	<b>SEPGGMQQAQ</b>	<b>SYRDAQ</b> TLVQ	<b>SDYTLQ</b> SNGH	50
51	PLSHAHQWIT	AL <b>SDG</b> PAW	<b>ASSPLGQQDI</b>	<b>KPSVQSSRDE</b>	<b>LHGAGTLQHQ</b>	100
101	<b>GRAPHLVHPA</b>	<b>HGNHGGP</b> AW	<b>RSTGSAHLS</b>	<b>MASSNGQGLL</b>	<b>YSQPSFTVNG</b>	150
151	<b>MINPGSGQGM</b>	<b>HHHGLRDAHD</b>	<b>DHHGEHGHQP</b>	<b>PPQTQ</b> QHSQ	<b>LQGGHHDSD</b>	200
201	<b>EDTPTS</b> DDLE	QFAKQFKQRR	IKLGFQADV	GLALGTLYGN	VFSQTTICRF	250
251	<b>EALQLSFKNM</b>	<b>CKLKPLL</b> NKW	LEEAD <b>SSSSGS</b>	<b>PTSIDKIAAQ</b>	<b>GRKRKRKRTS</b> I	300
301	<b>EVSVKGALES</b>	<b>HFLKCP</b> PSA	QEITSLADSL	QLEKEVVRVW	FCNRRQKEKR	350
351	<b>MT</b> PPGGTIPG	<b>PEDVYGASRD</b>	<b>TPPHGVQTS</b>	<b>VQ</b>		400