

Supplementary Figures

Role of *ZFHX4* in orofacial clefting based on human genetic data and zebrafish models

Nina Ishorst^{*‡} (1,2), Selina Hölzel^{*} (1,3), Carola Greve (1,#), Öznur Yilmaz (2), Tobias Lindenberg (2), Jessica Lambertz (2), Dmitriy Drichel (4), Berina Zametica (1), Enrico Mingardo (3), Jeshurun C. Kalanithy (1,2), Khadija Channab (3), Duygu Baydemir (1,2), Sabrina Henne (1), Franziska Degenhardt (1,†), Anna Siewert (1), Michael Dixon (5), Teresa Kruse (6), Edwin Ongkosuwito (7), Katta M. Girisha (8), Shruti Pande (8), Stefanie Nowak (1), Gregor Hagelueken (9), Matthias Geyer (9), Carine Carels (10), Iris A.L.M. van Rooij (11), Kerstin U. Ludwig (1), Benjamin Odermatt^{**‡} (2,3), Elisabeth Mangold^{**‡} (1)

Affiliations:

1 Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany

2 Institute of Anatomy, Division of Neuroanatomy, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany

3 Institute of Anatomy and Cell Biology, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany

4 Cologne Center for Genomics, University of Cologne, Cologne, Germany

5 Faculty of Biology, Medicine & Health, University of Manchester, Manchester M13 9PL, UK

6 University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Orthodontics, Cologne, Germany

7 Department of Dentistry, Section of Orthodontics and Craniofacial Biology, Radboud Institute for Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

8 Department of Medical Genetics, Kasturba Medical College, Manipal, Manipal Academy of Higher Education, Manipal, India

9 Institute of Structural Biology, University of Bonn, Bonn, Germany

10 Department of Human Genetics, KU Leuven, Leuven, Belgium

11 Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands

Current address: LOEWE Centre for Translational Biodiversity Genomics, Frankfurt am Main, Germany

† Current address: Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital Essen, University of Duisburg-Essen, Duisburg, Germany

* First authors contributed equally.

** Last authors contributed equally.

‡Corresponding authors:

Nina Ishorst: nina.ishorst@uni-bonn.de

Elisabeth Mangold: e.mangold@uni-bonn.de

Benjamin Odermatt: b.odermatt@uni-bonn.de

	hs_ZFHX4 (Q86UP3-5)	mm_ZFHX4 (H3BLK8)	dr_Zfhx4 (F6NXU9)
hs_ZFHX4 (Q86UP3-5)	100.00	92.17	72.23
mm_ZFHX4 (H3BLK8)	92.17	100.00	71.11
dr_Zfhx4 (F6NXU9)	72.23	71.11	100.00

hs_ZFXH4	METCDSPPISRQENGQSTSKLCGTTQLDNEVPEKVGAMFEDRENSSTDDNLKTDERKSEA	1
nm_ZFXH4	METCDSPPISRQENGQSTSKLCGTTQLDNEVPEKVGAGIEPDRENSSSHNLKTDERKSEV	60
dr_Zfxh4	METCDSPPLLRQENGQQISKIRETHLDNEVEKVPMEFPDKENSPPADNLRTEESLREI	60
	***** : ***** . *** * : ***** * : * : * : * : * : *	
hs_ZFXH4	LLGFSVENAATQVTSAKEIPCNCECATSFPSLQKYMHEHCNPNARLPVLKDDNESEISELE	120
nm_ZFXH4	LLGFSIENAAATQVTSAKEIPCNCECATSFPSLQKYMHEHCNPNARLPVLKDD-ESETSELE	119
dr_Zfxh4	ASGLSAENAI G---TATKEIPCNCECATSFSSLQKYMHEHCNPNARLPVLKDDNESEISDLE	117
	* . * *** . * : ***** ***** : * : * : * : *	
hs_ZFXH4	DSDEVNLTGEIVYQPDGSAYIIEDSKESGQNAQTGANSKLFSTAMFLDSLASAGEKSDQS	180
nm_ZFXH4	DSDEVNLTGEIVYQPDGSAYIIEDSKESGQNAQTGANSKLFSTAMFLDSLASAGEKSDQS	179
dr_Zfxh4	DSDEVNLTGEIVYQPDGSAFILEDKESGGQNAQSGVK-SLFSPALFSNSQTAAGDKTEQS	17
	***** : * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	ASAPMSFYPIINTFHIASSLGKPFADQAFPNTSALAGVGPVLHSFRVYDLRHKREKDY	240
nm_ZFXH4	STAPVSYFYPIINTFHIASSLGKPFADQAFPNTSALAGVGPVLHSFRVYDLRHKREKDY	239
dr_Zfxh4	ATAPMSFYPIQVINTFHIASSLGKPFADQAFPNTSALAGEGPVLHSFRVYDLRHKREKDY	236
	: * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	LTSDGSAKNSCVSKDVPNNVDLSKFDGCVSDGKRKPVLMCFLCKLSFGYIRS FVTHAVHD	300
nm_ZFXH4	LTSDGSAKNSCVSKDVPNNVDLSKFDGCVSDGKRKPVLMCFLCKLSFGYIRS FVTHAVHD	299
dr_Zfxh4	LITIDGAAKNSCVSKDVPNNVDLSKFDGCVSDGKRKPILMCFLCKLSFGYSRS FVTHAVHD	296
	** * : ***** : * : * : * : * : * : * : * : *	
hs_ZFXH4	HRMTLNDEEQKLNSKCVSAIIQGIQKDKPELISFLEPKKSTSVYPHFSTTNLIGDPDTF	360
nm_ZFXH4	HRMTLNDEEQRLNSKCVSAIIQGIQKDKPELISFLEPKKSTSVYPHFSTTNLIGDPDTF	359
dr_Zfxh4	HRMTLNDEEQKLNSKCVSAIIQGIQKDKPELISFLEPKKNSMLPFHFSTANFMGDPDGI	356
	***** : * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	RGLWSAFHVENGDLSLQAGFAFLKGSASTSSSAEQPLGITQMPKAEVNLGGLSSLVNTPI	420
nm_ZFXH4	RGLWSAFHVENGDLSLQAGFAFLKGSASPSSSAEQPLGITQMPKAEVNLGGLSSLVNTPI	419
dr_Zfxh4	RSLLNTHFMEGNDLSLQAGFAFLKGSASTSSSAEQPLGITQMPKAEVNLGGLSSLVNTPI	416
	* . * . * : * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	TSVSLSHSSSESSKMSKSKDQENNCERPKESNVLPNGECVPKSEPTPEGDEDEEDAYSIN	480
nm_ZFXH4	TSVSLSHSSSESSKMSKSKDQENNCERPKESTILHPNVGCPVKSEPTPEGDEDEEDAYSIN	479
dr_Zfxh4	SSVSLQRSSPG---AGYRDQESNCERQKDLSTLHANGELPIKSEPRDVADEEDMYSIN	472
	: * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	ELDDEEVLGELTDSIGNKDFPLLNQSIPLSSSVLKFIKGTSSSSATVSDDETEKKQTA	540
nm_ZFXH4	ELDDEEVLGELTDSIGNKDFPLLNQSIPLSSSVLKFIKGTSSSSATVSDDETEKKQAA	539
dr_Zfxh4	ELDDDGVC-CELEDSTSSKDFPLLNQSIPLSSSVLKFTERTGTSSVTV-ANDEFEKTKHAA	530
	*** : * * * * . : ***** * : * : * : * : * : *	
hs_ZFXH4	-AVRASGSVAANYISGKDFADASASKDSATAAHPSEIARGDEDSSATPHQHGFTPTSPG	599
nm_ZFXH4	AAGRSNGNVNTSYISGKDFADGSIISRDGTTA-APSETTHGDSDS-STTHQHGFTPTSPG	597
dr_Zfxh4	ANLDSYNDST---GGSKDCSDISGRRGSPSPLDLMRRDDESPGPLHGHAAATPTSPG	587
	. . . : . * * . * . : . : * : * : * : * : * : *	
hs_ZFXH4	TPGPGGDGSPGSGIECPKCDTVLSSSRSLGGHMTMMHSRNSCKTLKCPKCNWBYKYQQT	659
nm_ZFXH4	TPGPGGDGSPGNGIECPKCDTVLSSSRSLGGHMTMMHSRNSCKTLKCPKCNWBYKYQQT	657
dr_Zfxh4	TPGPG-EGSPGSGIECPKCDTVLSSSRSLGGHMTMMHSRNSCKTLKCPKCNWBYKYQQT	646
	***** : * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	EAHMKKEKHPGGSCVYCKTGQPHRLARGESYTCGYKPFRCVNCYSTTTKGNLSIHM	719
nm_ZFXH4	EAHMKKEKHPGGSCVYCKTGQPHRLARGESYTCGYKPFRCVNCYSTTTKGNLSIHM	717
dr_Zfxh4	DAHMKKEKHPGGSCVYCRTGQPHRLARGESYTCGYKPFRCVNCYSTTTKGNLSIHM	706
	: ***** : ***** : ***** : ***** : *****	
hs_ZFXH4	SDKHLNNVQNLQNGNGEVQFGHSAPAPNTSLSGCGTPSPSKPKQKPTWRCEVCDYETNVA	779
nm_ZFXH4	SDKHLNNVQNLQNGNGEVQFGHSAPNTSLSGCGTPSPSKPKQKPTWRCEVCDYETNVA	777
dr_Zfxh4	SDKHLNNVQTLQNGGTEAIYNHTTPTVTNTSLTGCCTGTPSPSKPKQKPTWRCEVCDYETNVA	766
	***** : * : * : * : * : * : * : * : * : * : *	
hs_ZFXH4	RNLRIHMTSEKHMHNMMMLLQQNMKQIQHNHLGLAPAEAEIYQYYLAQNIIGLTGMKLENP	839
nm_ZFXH4	RNLRIHMTSEKHMHNMMMLLQQNMKQIQHNHLGLAPAEAEIYQYYLAQNIIGLTGMKLENP	837
dr_Zfxh4	RNLRIHMTSEKHMHNMMMLLQQNMKQIQHNHLGLAPAEAEIYQYYLAQNIIGLTGMKLENP	826
	***** : ***** : ***** : ***** : *****	
hs_ZFXH4	ADPQLMINPFQLDPATAAALAPGLVNNELPPEIRLASGQLMGDDLSLTAGELSPYISDP	899
nm_ZFXH4	AETOLLNPFQFDSATAAALAPGLVNNELPPEIRLASGQLMGDDLSLTAGELSPYISDP	899

Girisha & colleagues
p.Tyr653His

dr_Zfhx4	SDPQMMINPFQLDPSTAAALTPGHVNNELPAELRLASGQLMGDDLSVLSAGELSPCINDP	886
hs_ZFHX4	ALKLFQCAVCNKFTSDSLEALS SVH SSERSLP EE EWRAVI GDIYQCKLCNYNTQLKANFQ	959
mm_ZFHX4	ALKLFQCAVCNKFTSDSLEALS SVH NSERSLP EE EWRAVI GDIYQCKLCNYNTQLKANFQ	957
dr_Zfhx4	LLKLFQCAVCNK FSS DSLEALS GH VATERSLP EE EWRAVMDVYQCKLCNYNTQLKANFQ	946
hs_ZFHX4	LHCKTDKHM QKYQLVAHIKEGGKSNEWR LKCIAIGNPVHLKCNACDY TNSVDKLR LH TT	1019
mm_ZFHX4	LHCKTDKHM QKYQLVAHIKEGGKSNEWR LKCIAIGNPVHLKCNACDY TNSVDKLR LH TT	1017
dr_Zfhx4	LHCKTDKHM QKYQLVAHIKEGGKANQWR LKCIAIGNPVHLKCNACDY SNSVEKLR LH AT	1006
hs_ZFHX4	NHRHEAALKLYKHL KQEGAVNPE SCY YYCAVCDYTTKVK LN LVQHVRSVKH QQT EGLRK	1079
mm_ZFHX4	NHRHEAALKLYKHL KQEGAVNSE SCY YYCAVCDYSSKIK LN LVQHVRSVKH QQT EGLRK	1077
dr_Zfhx4	NQRHEAALKVYKHL KQESACNSE SCY YYCALCDYSSRAK LN LLQHLRSVKH QQT SEGLRK	1066
hs_ZFHX4	LQLHQQGLAPEEDNLSEIFFVKDCPPNELETASLGARTCDDDLTEQQLRSTSEEQSEEA	1139
mm_ZFHX4	LQLHQQGLPSEEDNLSEIFFVKECPANELETASLGARNGEDELIEQQLKAASEEPSEDA	1137
dr_Zfhx4	LQLHQQGLPPDEDNLADIFLVKDCPPNESEEPGEDS-----E	1103
hs_ZFHX4	GAIKPTA VAEDDEK DTSERD NS EGKNSNKD SGIITPEKELKVS VAGGTQPLLAKEDVA	1199
mm_ZFHX4	DPLK PPTVAEDDEK EAHKRD NS EGKISTKDP-VIVPEKELKV-VT GATQPLLAKED NTG	1195
dr_Zfhx4	EF PQASSLAA EDKDSKRD AI E GKTT EKDSRINT LAKETLG IATAGKQSLQ QEKEND TP	1163
hs_ZFHX4	TKRSKPTEDNKFCE QFYQCPYCNYSRDQSR I QMHVLSQHS VQPVIC CP LQ QD VLSN KM	1259
mm_ZFHX4	TKRSKPTEDNKFCE QFYQCPYCNYSRDQSR I QMHVLSQHS VQPVIC CP LQ QD VLSN KM	1255
dr_Zfhx4	PKR PKSA E EK TL NS EQV QCPYCNYSNKDANRLQLH IMS QHS MQPVIS CP LQ QD VLSN KI	1223
hs_ZFHX4	HLQLHL THLH SVSPDCVEKL LMTVPVPDVMMPNSMLLPAA ASEK SERDTPAAVT AEGSGK	1319
mm_ZFHX4	HLQLHL THLH SVSPDCVEKL LMTVPVPDVMMPNSMLLPAA ASEK SERDTPAAVT AEGSGK	1315
dr_Zfhx4	HLQLHL THLH SVAPDCVEKL LMTVAGPDL PVPSSLLSASLQ QD KVPS LM DTSAVN PE GSGK	1283
hs_ZFHX4	YSG SPMD DKSMAGLED SKANVEV KNEEQKPTKEPLEVSEWNKNS SKDVKIP DTLQ QD LN	1379
mm_ZFHX4	YSG SPVD DKSMGL EDSKVGEIKNEEQKPAKEPVE ASEWNK TSKDVNI SDALQ QD LN	1375
dr_Zfhx4	TM G NSSKDLKNGT--GQDKSEV DLTSEELKPLKEAAE APDWK KA SGQDRKSP DALQ EHLS	1341
hs_ZFHX4	EQQKRQPLSVSDRHVYKYRCN HCSLAFKTMOKLQIHSQYHAIRAA TM CNLCQ RS FRTFQA	1439
mm_ZFHX4	EQQKRQPLSVSDRHVYKYRCN HCSLAFKTMOKLQIHSQYHAIRAA TM CNLCQ RS FRTFQA	1435
dr_Zfhx4	DLQKRQPLSVSDRHVYKYRCN HCSLAFKTMOKLEIHSQYHAIRAA TM CNLCQ RS FRTFLA	1401
hs_ZFHX4	LKKHLEAGHP ELSEAE LQQLYASLPVNGELWAESETMSQDDHGLEQEMEREYEV DHEGKA	1499
mm_ZFHX4	LKKHLEAGHP ELSEAE LQQLYASLPVNGELWAESETMTQDDHGIDQEMEREYEV DHEGKA	1495
dr_Zfhx4	LRKHLE TGHP ELTEAEVQQLCGNLP NGDI SESEMRAL EEAQA FENELDKDEEL DQEGKA	1461
hs_ZFHX4	SPVGS SSSI PD DMG SEPKRTL PFRKGNF TMEKFLDPSR PKCTVCKESFTQKN ILLVH	1559
mm_ZFHX4	SPV ESD SSSI PD DLGLEPKRTL PFRKGNF TMEKFLDPSR PKCTVCKESFTQKN ILLVH	1555
dr_Zfhx4	SPTGSDSSSLDDMGSEPKRTL PFRKGNF TMEKFLDPSR PKCTVCKESFTQKN ILLVH	1521
hs_ZFHX4	VNSVSHLHKLK KKVLQ EASSPV QETNSNTDNKPKYC IC NVAYSQS STLEI HMR SVL HQT	1619
mm_ZFHX4	VNSVSHLHKLK KKVLQ EASSPV QETNSNTDNKPKYC STCSVAYSQSSTLEI HMR SVL HQT	1615
dr_Zfhx4	VNSVSHLHKLK KKVIQ EASSPV QETNSNVDNPKYCN IC NVAYSQS STLEI HMR SVL HQT	1581
hs_ZFHX4	KARAAKLEPSGHVAGGHSIAAN--VNSPGQGM LDMS LA AV -NSK DTHLDAKE LNKKQT	1675
mm_ZFHX4	KARAAKLEPSRHLP SGHS ITAA--VNSPGQGM LESMS LA SV -NSK DTHLDAKE LNKKQT	1671
dr_Zfhx4	KARTAKLETSTSTGSGSSGSTSSKSPVPPNHGNTDSASAAPVSTNKENTVDAKEVT MKTQ	1641
hs_ZFHX4	PDLISAQPAHHP-PQSPAQIQMLQHE LQQQA AF FQ QFLNPAFLPHF FMTPEALLQ FQ	1734
mm_ZFHX4	PDLISAQPTTHPPPRSPAQIQMLQHE LQQQA AF FQ QFLNPAFLPHF FMTPEALLQ FQ	1731
dr_Zfhx4	TEHIPVQPN-HPPTQSPAQLQMLQHE LQQQA AF FQ QFLNPAFFPHF FMTPEALLQ FQ	1700
hs_ZFHX4	PQFLFPFYIPGTEFSLGPD LGLPGSATFGMPGMTMAGS LLEDLKQ QIQ TQH HV GQTQ LQ	1794
mm_ZFHX4	PQFLFPFYIPGA EFSLGPD LGLPTSTTFGVPGMTMAGS LLEDLKQ Q QIQ TQH HV GQTQ LQ	1791
dr_Zfhx4	PQFLFPFYIPGA EFNIS PELALHS-AAFGMPG---MTGSFLEDLKQ QMQQ QHLQ QQQ Q	1756
hs_ZFHX4	ILQQQAQQYQATQPQLQ PQKQQQ PPPPQQQQQQQASKLLKQE QSN IVSAD CQIMK DVPS	1854
mm_ZFHX4	FLO-QAQQYQAVQPQLQ PQNOQ -PPLPQQQQPQQPSKLLKQE QSL ASTDCQLMK DMPS	1849
dr_Zfhx4	QLSQQAQQQASQSQMQQQK VQQQSHKPKTES -----HNVLSEIQ MSRD ---	1801
hs_ZFHX4	YKEAEDISEKPEKPKQEFISEGEG LKEGKDTKKQKSLE PSIPPPRIASGARGNAAKALLE	1914
mm_ZFHX4	YKEAEEVTEKQEKPKQEFINDTEGLKDSKDKKQKSLEPCIPPPRIASGARGNAAKALLE	1909
dr_Zfhx4	--AEEHLEKQEGKAKQDTANENDNGKDAKDN RKPFSE PLIPPPRIISGARGNAAKALLE	1859
hs_ZFHX4	NFGFELVIQYNENRQKVQKKGKSGEGENTDKLECGT CGKLF SNVLI LKSHQ EHVHGQFFP	1974
mm_ZFHX4	NFGFELVIQYNENRQKVQKKGKSGEGENS DKLECGICGKLF SNVLI LKSHQ EHVHGQFFP	1969

Index Family BN00337
p.(Lys1035_Asp3612del
insAsn)

Family BN00648
c.3964+2T>A
splice donor

dr_zfhx4	NFGFELVIQYNENRQKSQKKNK--EDELTDKLECGLCGLFSNMLILKSHQEHIGHQFFP	1917
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	YAALEKFARQYREAYDKLYPISPSSETPPPPPPPPLPPAPPQSSMGVQIPNTVSTP	2034
mm_ZFHX4	YGALAEKFARQYREAYDKLYPISPSSETPPPPPPPPLPPAPPQSTLGPVQIPNTVSAP	2029
dr_zfhx4	YVELEKFAQQYREAYDKLYPINPASSETPPPPPPPPPPPPAPLNTL-PANA--SVG-K	1973
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	LQAPPPTPPPPPPPPPP-PPPPPPPPSAPPQVQLPVSLDLPLFPSIMMQPVQHPALPP	2093
mm_ZFHX4	LQAPPPTPPPPPPPPPPPPPPPPPPSAPPQVQLPVSLDLPLFPSIMMQPVQHPALPP	2089
dr_zfhx4	SQTPSPAPI-QTPQQAQPPPPPPPPPTAPPQVQLPVSLDMPFIFFPLMMQSVQHPGLPP	2032
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	QLALQLPMDALSADLTQLCQQQLGLDPNFLRHSQFKRPRTRITDDQLKILRAYFDINNS	2153
mm_ZFHX4	QLALQLPMDTSLADLTQLCQQQLGIDPNFLRHSQFKRPRTRITDDQLKILRAYFDINNS	2149
dr_zfhx4	QLALQLPTMDSLSSDLTQLCQQQLGLDPNFLRHSQFKRPRTRITDDQLKILRAYFDINNS	2092
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	PSEEQIQEMAESGLSQKVIKHWFRNTLFKERQRNKDS	2213
mm_ZFHX4	PSEEQIQEMAESGLSQKVIKHWFRNTLFKERQRNKDS	2209
dr_zfhx4	PNEEQIQEMADKSLPQKVIKHWFRNTLFKERQRNKDS	2152
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	SLEHYKSDASFSCRSSRTRETDYQLRVLQDFDNTNAYPKDDEIEQLSTVLNLPTRVIVVW	2273
mm_ZFHX4	SLEHYKSDAASFSCRSSRTRETDYQLRVLQDFDNTNAYPKDDEIEQLSTVLNLPTRVIVVW	2269
dr_zfhx4	LLEYTRIDSAANKSSRTRETDYQLRVLQDFDNTNAYPKDDEIEQLSTVLNLPTRVIVVW	2212
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	FQNAQKARKSYENQAETKDNEKRELTNERYIRTSNMQYQCKKCNVFPRIFDLITHQKK	2333
mm_ZFHX4	FQNAQKARKSYENQAETKDNEKRELTNERYIRTSNMQYQCKKCNVFPRIFDLITHQKK	2329
dr_zfhx4	FQNAQKARKTYENQADTKSEKKELTNERYIRTSNMQYQCKKCSVIFPRIFDLITHQKK	2272
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	QCYKDEDDDAQDESQTEDSDMATDQVVKHCTVSGQTDAAKNAAPAASSGSGTSTPLIP	2393
mm_ZFHX4	QCYKDEDDDAQDESQTEDSDMATDQVLYKCHMVSGQTDATKATLVAASSGSGTSTPLIP	2389
dr_zfhx4	QCYKDEDDDAQDEHILEDLSENIEQSLAKPME-----AAKQ--SFVTASSGSGSSPLMP	2324
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	SPKPEPEKTSKPPEYPAEKPKQSDSPSPSQGTKPALPLASTSSDPQASTAQPPQPQPP	2453
mm_ZFHX4	SPKPEPEKNSPKTEYPGEKTKQSDPSLPQG-TKSPSSVLTSSPQQA--SIPQPTQPP	2446
dr_zfhx4	SPRPDIKTSKPPELLNEKSKLGETAPLQTPKNPNEL-----KPSKAS--TPQPLSQKV	2376
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	KQPQLIGRPPSASQTPVPSSPLQISMTSLQNSLPPQLLYQCDQCTVAFPTLELWQEHQH	2513
mm_ZFHX4	KQPQLIGRPPSASQTPVPSSPLQISMTSLQNSLPPQLLYQPCDQCTIAFPTELEWKEHQH	2506
dr_zfhx4	FQPQLSRPHSQPQSTVPVSSPLSLALSSLNNSLPPQLQYHCEQCKIAFPTELEWQEHQH	2436
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	MHFLAAQNQLHSPFLERPMDMPYMIFDPNNPLMTGQLLGSSLTQMPPQASSSHTTAPT	2573
mm_ZFHX4	MHFLAAQNQLHSPFLERPMDMPYMIFDPNNPLMTGQLLGSSLTQMPPQTSTAHTTAPAS	2566
dr_zfhx4	MHFLAAQNQLHSPFLERPMDMPYMIFDPNNPLMTGQLLGSSLTQMPPQTSTAHTTAPAS	2492
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	VAASLKRKLDDKEDNNCSEKEGGNSGEDQHRDKRLRTTITPEQLEILYKYLDSNPTRK	2633
mm_ZFHX4	VAASLKRKLDDKEDNNCSEKEGGNSGEDQHRDKRLRTTITPEQLEILYKYLDSNPTRK	2626
dr_zfhx4	NSNSLKRKLDEKEDG-SHEKGLNSAEDQHRDKRLRTTITPEQLEILYKYLDSNPTRK	2551
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	MLDHIAREVGLKKRVVQVWFQNTARERKGOFRAGVPAQSHKRCPCRALFKAKSALESH	2693
mm_ZFHX4	MLDHIAREVGLKKRVVQVWFQNTARERKGOFRAGVPAQSHKRCPCRALFKAKSALESH	2686
dr_zfhx4	MLDHIAREVGLKKRVVQVWFQNTARERKGOFRAGVPSQTHKKPCPCRALFKAKSALDSH	2611
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	IRSRHWNEGKQAGYSLPPSPLISTEDGGESPKYIYFDYPSLPLTKIDLSSENEASTVS	2753
mm_ZFHX4	IRSRHWNEGKQAGYSLPPSPLISTEDGGESPKYIYFDYPSLPLTKIDLSSENEASTVS	2746
dr_zfhx4	IRSRHWHEAKQAGFSLPPSPMPQDDEGQSPHKYSFSDYLQMPPTKI-EMN-ENEPTASS	2669
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	TPVSKTAELS PKNLLSPSSFKAECSEDEVENLNAPPAEAGYDQNKTFDETSSINTAISDA	2813
mm_ZFHX4	TPVSKTAELS PKNLLSPSSFKAECSEDEVENLNAPPAEAGYDQSKTFDETSSINTAISDA	2806
dr_zfhx4	TPVK-PSETLTKNPLNIPSLKAHSDELEALNMSSAEASFDTNKMDDETSSINTAVSDA	2728
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	TTGDEGNTMESTTGSSGDVVKPALSPKEPKTLDTLPKPATPTTEVCDDKFLFLTSPSI	2873
mm_ZFHX4	TTGDEGAADMENT-GGSGEVKPALSPKETKTLDSLQKPATPTTEVCDDKFLFLTSPSI	2865
dr_zfhx4	TTGDETNEVENLTINGSEKM-----I--ENKSNQAQTS DMNEDRPFPMVSPAL	2776
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	HFNDKGDGHDQSFYITDDPDDNADRSETSSIADPSSPNPFGSSNPFKSK--SNDRPGHKK	2931
mm_ZFHX4	HFNDKGDGHDQSFYITDDPDDNADRSETSSIADPSSPNPFGSSNPFKSK--SNDRPGHKK	2923
dr_zfhx4	SFSGKSDS--YFNRYRDDFDNADRSETSSLADPSSPNPFGSTNPFKSSKAGGERPGHKK	2834
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	FRTQMSNLQLKVLKACFSDYRTPTMQCEMLGNEIGLPKRVVQVWFQNAKAKKFKINI	2991
mm_ZFHX4	FRTQMSNLQLKVLKACFSDYRTPTMQCEMLGNEIGLPKRVVQVWFQNAKAKKFKINI	2983
dr_zfhx4	FRTQMSNLQLKVLKACFSDYRTPTMQCEMLGNEIGLPKRVVQVWFQNAKAKKFKINI	2894
hs_ZFHX4	*****	
mm_ZFHX4	*****	
dr_zfhx4	*****	
hs_ZFHX4	GKPFMINQGGTGKPECTLCGVKYSARLSIRDHIFSKQHISKVRETVGSQLDREKDYLA	3051
mm_ZFHX4	GKPFMINQGGTGKPECTLCGVKYSARLSIRDHIFSKQHISKVRETVGSQLDREKDYLA	3043

dr_Zfhx4	GKPFMISQGS PDGPRPECTLCGVKYTARLSIRDHIFSKQHIAKVQETLGNQVDREKDYLA	2954
hs_ZFHX4	PTTVRQLMAQQELDRIKKASDVLGLTVQQPGMMDSSSLHGISLPTAYPGLPGLPPVLLPG	3111
mm_ZFHX4	PTTVRQLMAQQELDRIKKASDVLGLTVQQGITDNCSLHGISLQAAYPGLPGLPPVILPG	3103
dr_Zfhx4	PTTVRQLMAQQELDRLLKATDVLNLPAQQQPAVDNNALHGLSLPTAYPGISGLPPVLLPG	3014
hs_ZFHX4	MNGPSSLPFGFPQNSNTLTTPPGAGMLGFPTSATSSPALSLSAPTCKPLLQTPPPPPPPPPP	3171
mm_ZFHX4	MNGPSSLPFGFPQNSNTLTSPGTGMLGFPSATSSPALSLSGPTKSLQTPPPPPPPPP--	3161
dr_Zfhx4	VNGPSSLPFGFPNTPALASPGAGMLGFPTPATPSPAMSLSSTPTKLLQTPPPPPPPQPS	3074
hs_ZFHX4	---PPSSSLSGQQTQQNKESEKKQTKPNKVKKIKEEELEATKPEK-HPKKEEKISSALS	3227
mm_ZFHX4	---PPSSSLSGQQTQQNKESEKKQTKPNKVKKIKEESEAIKPEK-HPKKEEKISSALT	3217
dr_Zfhx4	ALPLPPTPLAANQTEQHIKESEKDKKL--EKPKVKEREADTARPETPIVKKKEKPSPLS	3132
hs_ZFHX4	VLGKVVGETHVDPIQLQALQNAIAGDPASFIGGQFLPYFIPGFASYFTPLPGTVQGGYF	3287
mm_ZFHX4	VLGKVVGETHMDPTQLQALQNAIAGDPASFIGGQFLPYFIPGFASYFSPQLPGTVQGGYL	3277
dr_Zfhx4	MLGKLGSEGGMDPAQLQALQNAIAGDPGSFLGGQFLPYFIPGFASCFSPQLPGGVQGGYF	3192
hs_ZFHX4	PPVCGMESLFYPGTPMPQTLAAGLSPGALLQQYQQYQQNLQESLQKQKQQQEQQKPVQA	3347
mm_ZFHX4	PPICGMESLFYPGPAVPQTLAAGLSPGALLQQYQQYQQSLQDSLQKQKQQQEQQKPVFA	3337
dr_Zfhx4	PPLCGMENLFYPGPAMPQAIAGLSPGALLQQYQQYQQSLQDSLQKQKQQQEQQKQSEFQ	3252
hs_ZFHX4	KT--SKV---ESDQPQNSNDAETKEDKSTATESTKEEPQLESKSAFSDTYVVPFVKY	3401
mm_ZFHX4	KT--AKG---EGDQPQSSNEAETKEEKSTAPESTKEEVQLDSKSAFSDTCIVPFVKY	3391
dr_Zfhx4	RKTPAPMKSTSAQNSNLKPKETVETKDDKGSSTESTNEEPQNTNKSSGFPDAFIVPSIKH	3312
hs_ZFHX4	EFICRKQCMFTDEDAAVNHQKSFICYFGQPLIDPQETVLRVPVSKYQCLACDVASGNEA	3461
mm_ZFHX4	EFVCRKQCMFTDEDAVNHQKSFICYFGQPLIDPQETVLRIPVSKYQCLACDLASGNEA	3451
dr_Zfhx4	EFICRKQCMIFAEDSAARHQKSFICYFGHPFIDPQETVLRKAVSKYNCIACNVSVSGNEA	3372
hs_ZFHX4	LSQHLQSSSLHKEKTIKQAMRNAKEHVRLLPHSVCSNPNTTSTSQAASSNNTYPHLSCF	3521
mm_ZFHX4	LSQHLQSSSLHKEKTIKQAMRNAKEHVRLLPHSVCSPPNTSSTSPSAASSNNTYPHLSCF	3511
dr_Zfhx4	LGQHLQSSSLHKEKTIKQAMRNAKEHARLLPHSVCSNPNTTSTSQAASSNNTLPHLSHL	3432
hs_ZFHX4	SMKSWPNILFQASARRAAS-PPSSPPSLSLPSTVTSSLCSTSGVQTSPLTESCSDSDSE	3580
mm_ZFHX4	SMKSWPNILFQASARKAAS-SPSSPPSLSLPSTVTSSLCSTSGVQTSPLTESCSDSDSE	3570
dr_Zfhx4	SMKSWPNILFQASARKAASCPASPPPLSLPSTVTSTSCSTSGVQTSPLTESCSDSDSE	3492
hs_ZFHX4	LSQKLEDLDNSLEVKAKPASGLDGNFNSIRMDMFSV	3616
mm_ZFHX4	LSQKLQDLDNSLEVKAKPASGLDGNFNSVRMDMFSV	3606
dr_Zfhx4	LSQKLDDLDALEAKAKAASGLDGNFSSMRMDMFSV	3528

Figure S1. Conservation of ZFHX4 on amino acid level created with Clustal Omega (Madeira et al., 2022). (A) ZFHX4 amino acid sequence identity between human (hs) mouse (mm) and zebrafish (dr) in percent. UniProt sequence identifier given in brackets. **(B)** ZFHX4 protein sequence alignment (hs, mm and dr). Amino acid residues are colored corresponding to their physiochemical features: red - small, hydrophobic (including aromatic, without Y); blue - acidic; purple - basic (without H); green - hydroxyl, sulfhydryl, amine and G. Asterisk, full conservation in that position; colon or period, conservation between residue groups with strongly (:) or weakly (.) similar features. Positions of zinc finger (grey) and homeobox (yellow) domains are highlighted in hs ZFHX4 sequence. Variant locations from present study, Zametica et al. and Girisha K.M. & colleagues are marked.

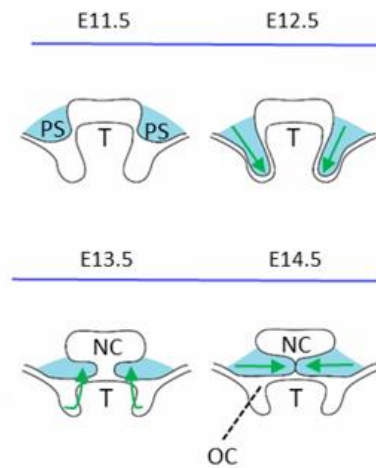


Figure S2. Timeline of secondary palate development in mice (E11.5–E14.5). Anatomical structures that were dissected for the bulk RNA-Seq dataset (Figure 2A) are given in light blue. PS, palatal shelves; T, tongue; NC, nasal cavity; OC, oral cavity.

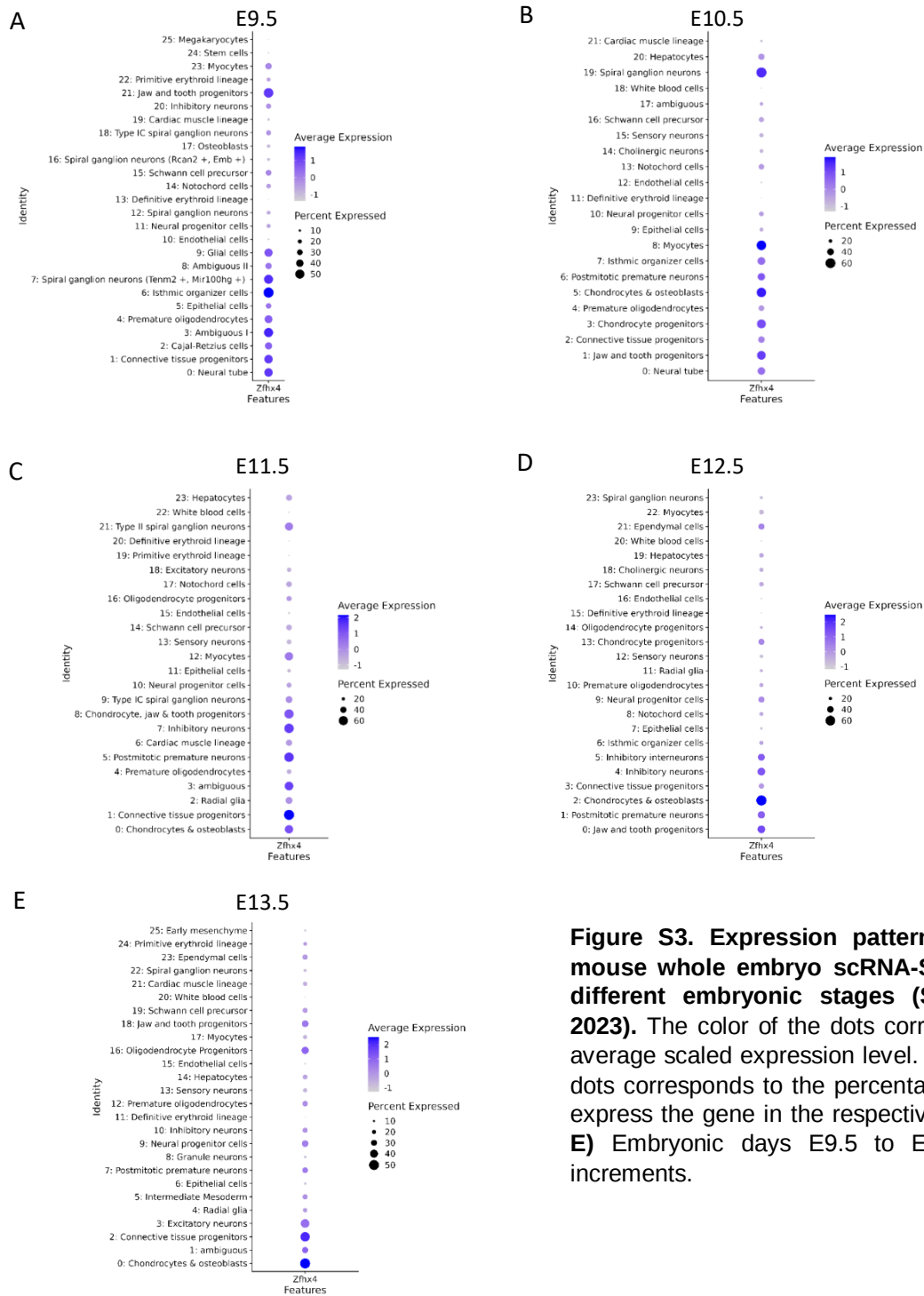
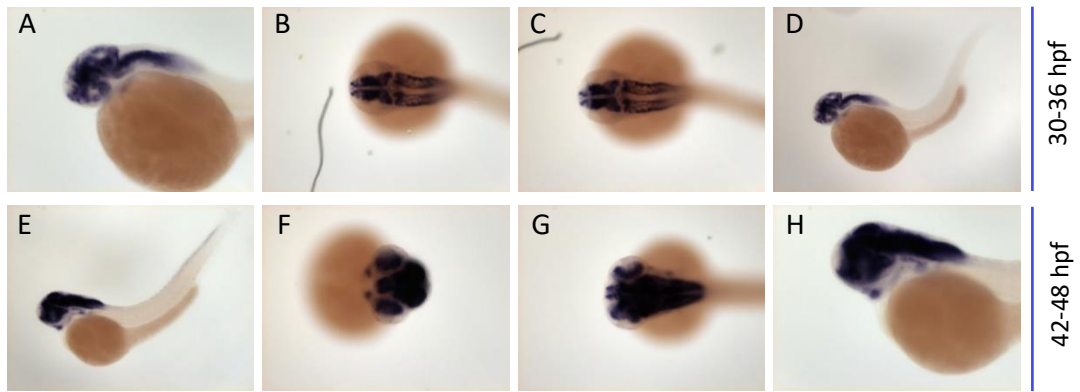
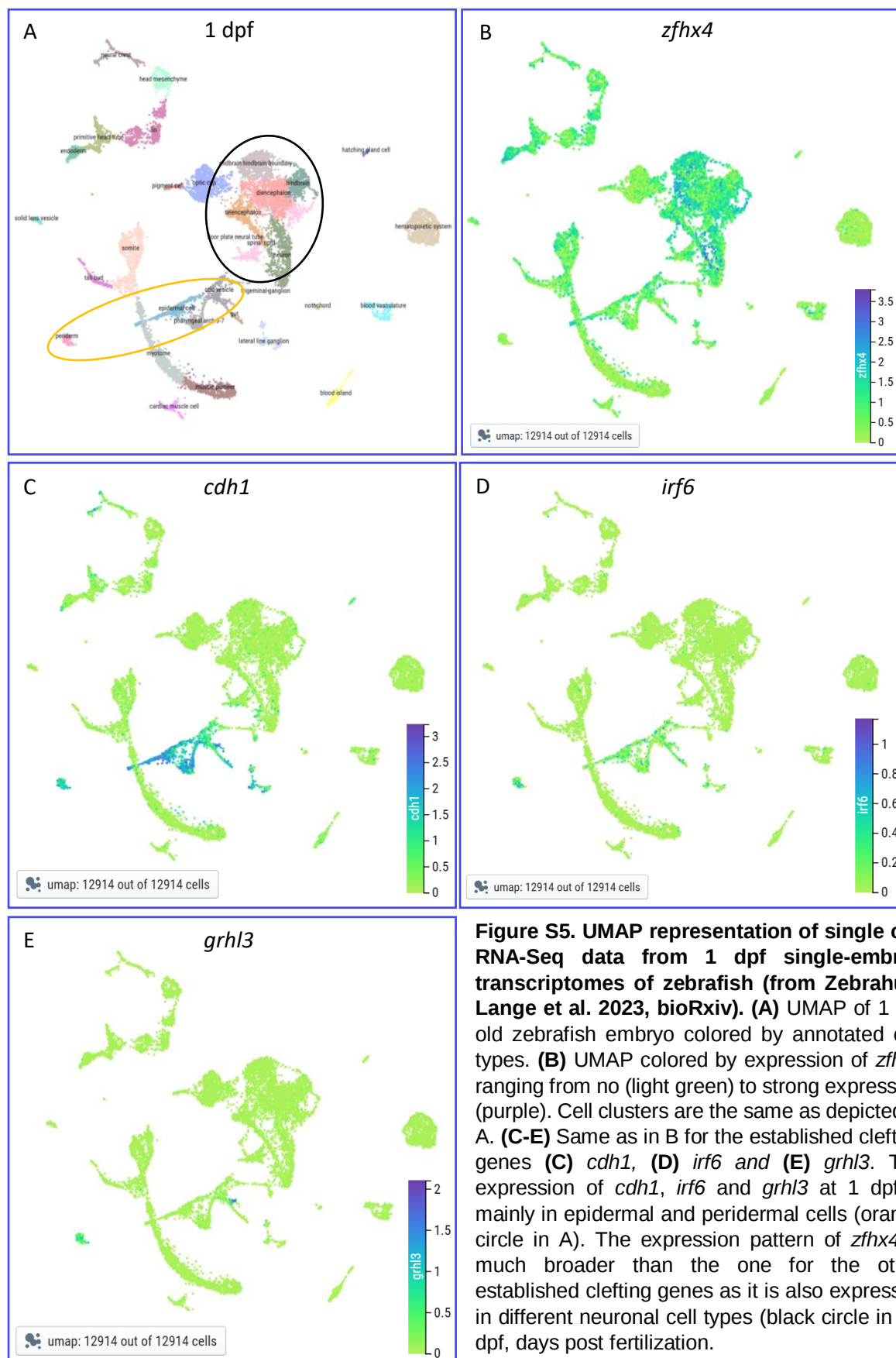


Figure S3. Expression pattern of *Zfhx4* in mouse whole embryo scRNA-Seq data from different embryonic stages (Siewert et al. 2023). The color of the dots corresponds to the average scaled expression level. The size of the dots corresponds to the percentage of cells that express the gene in the respective cell type. **(A-E)** Embryonic days E9.5 to E13.5 in 1-day increments.



Thisse & Thisse 2004

Figure S4. Whole mount *in situ* hybridization of *zfhx4* in zebrafish larvae (zfl) of different developmental stages from Thisse, B., Thisse, C. (2004) Fast Release Clones: A High Throughput Expression Analysis. ZFIN Direct Data Submission. (<http://zfin.org>). (A-D) Zfl from Prim 15 – Prim 25 stage (30-36 hpf). Expression according to Thisse, B., Thisse, C., 2004 annotations in diencephalon, hindbrain, midbrain and telencephalon as indicated by blue/purple dye. (E-H) Zfl from High-Pec – Long Pec stage (42-48 hpf). Expression according to Thisse, B., Thisse, C., 2004 annotations in brain, pharyngeal arch and retinal ganglion cell layer as indicated by blue/purple dye. Zfl, Zebrafish larvae; hpf, hours post-fertilization.



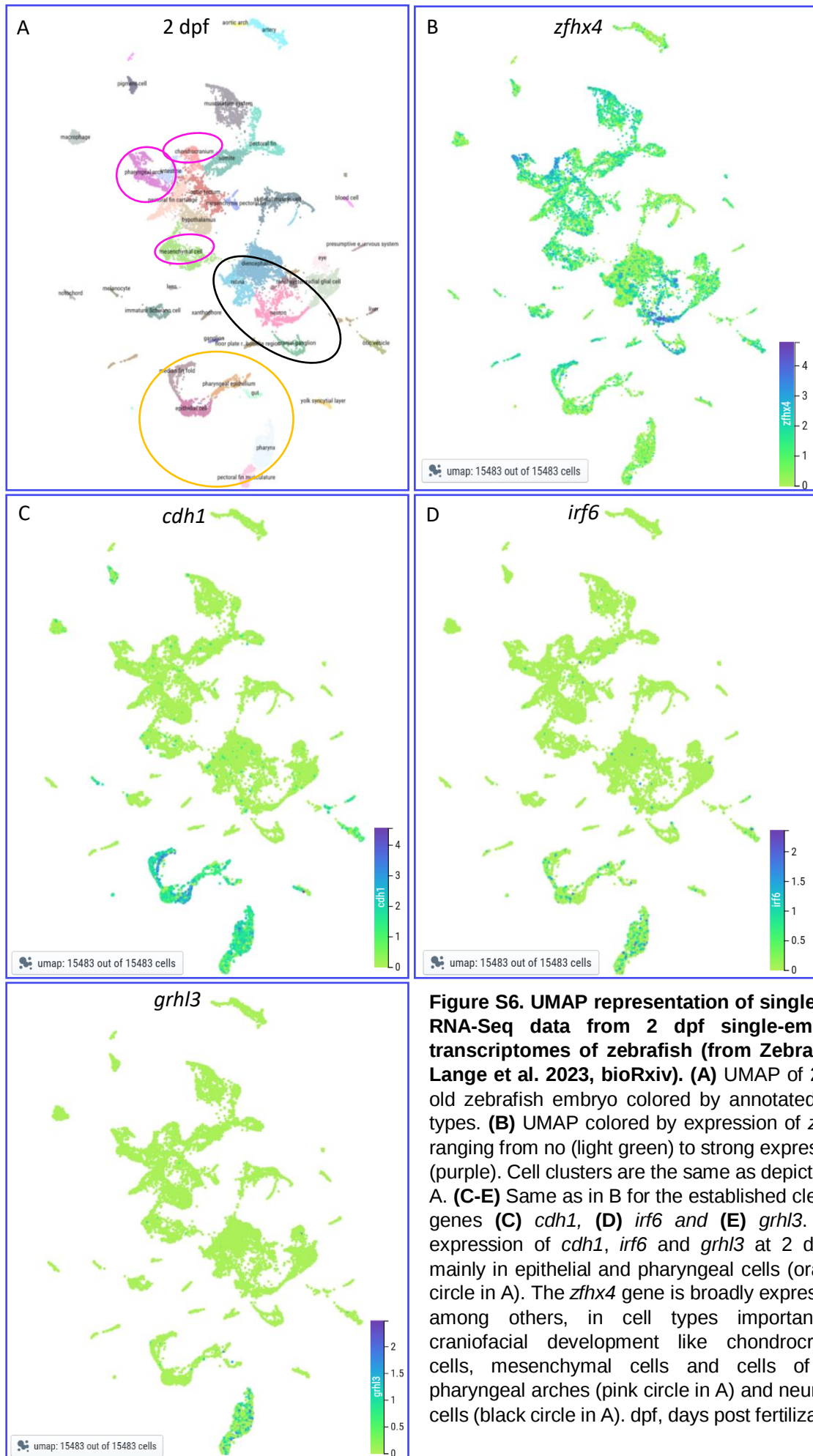
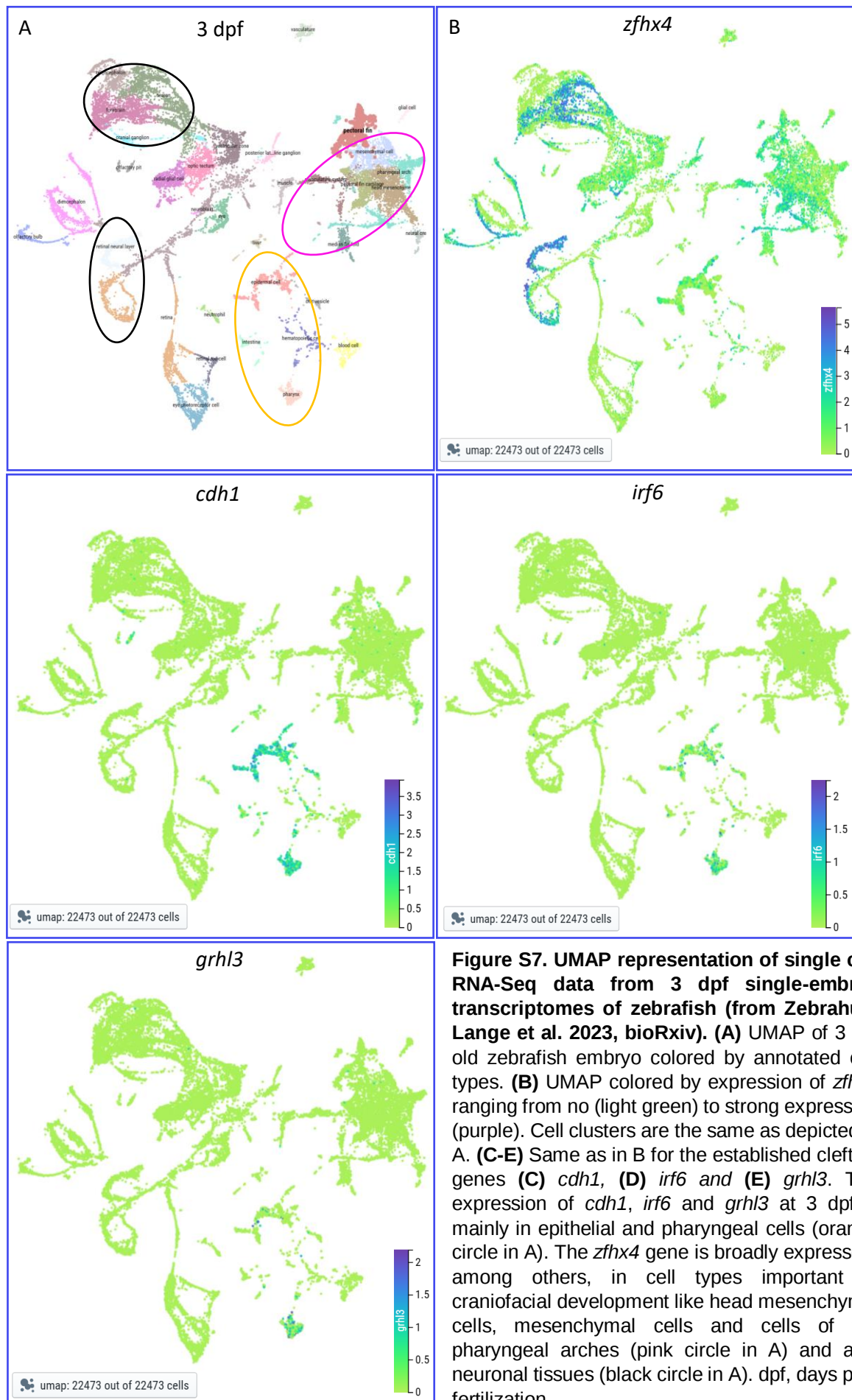


Figure S6. UMAP representation of single cell RNA-Seq data from 2 dpf single-embryo transcriptomes of zebrafish (from Zebrafish, Lange et al. 2023, bioRxiv). (A) UMAP of 2 dpf old zebrafish embryo colored by annotated cell types. (B) UMAP colored by expression of *zfhx4* ranging from no (light green) to strong expression (purple). Cell clusters are the same as depicted in A. (C-E) Same as in B for the established clefting genes (C) *cdh1*, (D) *irf6* and (E) *grhl3*. The expression of *cdh1*, *irf6* and *grhl3* at 2 dpf is mainly in epithelial and pharyngeal cells (orange circle in A). The *zfhx4* gene is broadly expressed, among others, in cell types important to craniofacial development like chondrocranial cells, mesenchymal cells and cells of the pharyngeal arches (pink circle in A) and neuronal cells (black circle in A). dpf, days post fertilization.



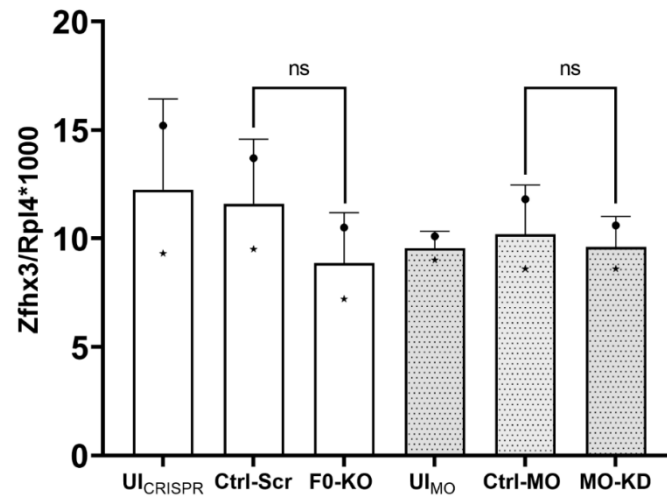


Figure S8. Detection of Zfhx3 with Liquid Chromatography Mass Spectrometry (LC-MS) analysis of protein lysates from heads of 3 dpf zfl (N=2) to show specificity of F0-KO and MO-KD. Depicted are Zfhx3 fold changes relative to 60S ribosomal protein L4 (Rpl4) abundance of each sample. Zfhx3 levels were not significantly decreased in F0-KO compared to Ctrl-Scr and in MO-KD compared to Ctrl-MO. Individual data of LC-MS run1 (stars) and run2 (dots) are plotted. Statistical testing was performed using two-way analysis of variance (ANOVA), Tukey's multiple comparisons test. Data are presented as means with standard deviation (SD). ns = not statistically significant. dpf, days post-fertilization; UI_{CRISPR/MO}, uninjected wildtype zfl; Ctrl-MO, Control MO injected zfl; Ctrl-Scr, Scrambled control injected zfl.