Supplementary Information:
Zebrafish as a model to study live mucus physiology

*Correspondence: ribbeck@mit.edu

Irena Jevtov[1], Tore Samuelsson[2], Grace Yao[1], Adam Amsterdam[3], Katharina Ribbeck[1]

[1]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA

[2]Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy at University of Gothenburg, Göteborg, Sweden

[3]David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA

Supplementary Information

Supplementary Figure S1. Genomic structure of five predicted polymeric secreted mucins in zebrafish.

Supplementary Figure S2. Neighbor-joining phylogenetic tree of mucins in zebrafish, human, and chicken.

Supplementary Figure S3. Alignment of gel-forming secreted mucins from zebrafish, human, and chicken.

Supplementary Figure S4. *In situ* hybridization of the secreted mucins in sagittal sections of adult zebrafish.

Supplementary Figure S5. Differential expression of *muc5.1 and muc5.2* in gill lamellae.

Supplementary Figure S6. qRT-PCR analysis of mucin gene expression during zebrafish development.

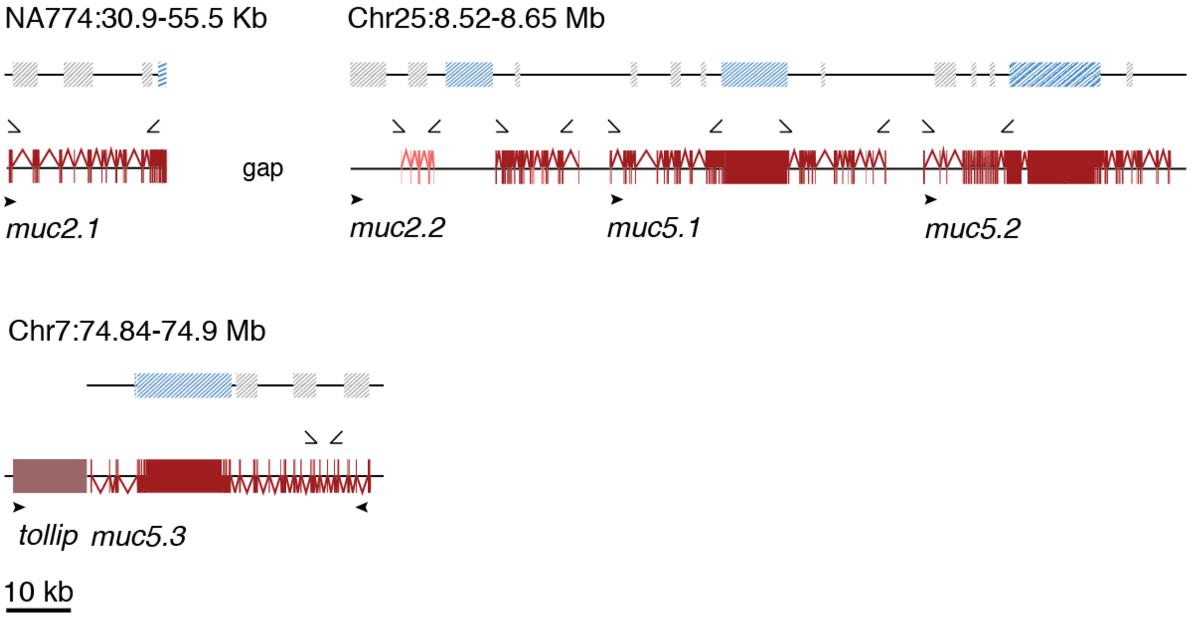
Supplementary Figure S7. Generation of the fluorescent mucin reporter *muc5.1:S-RFP*.

Supplementary Table S1. Additional identified mucin-like fragments in the zebrafish genome.

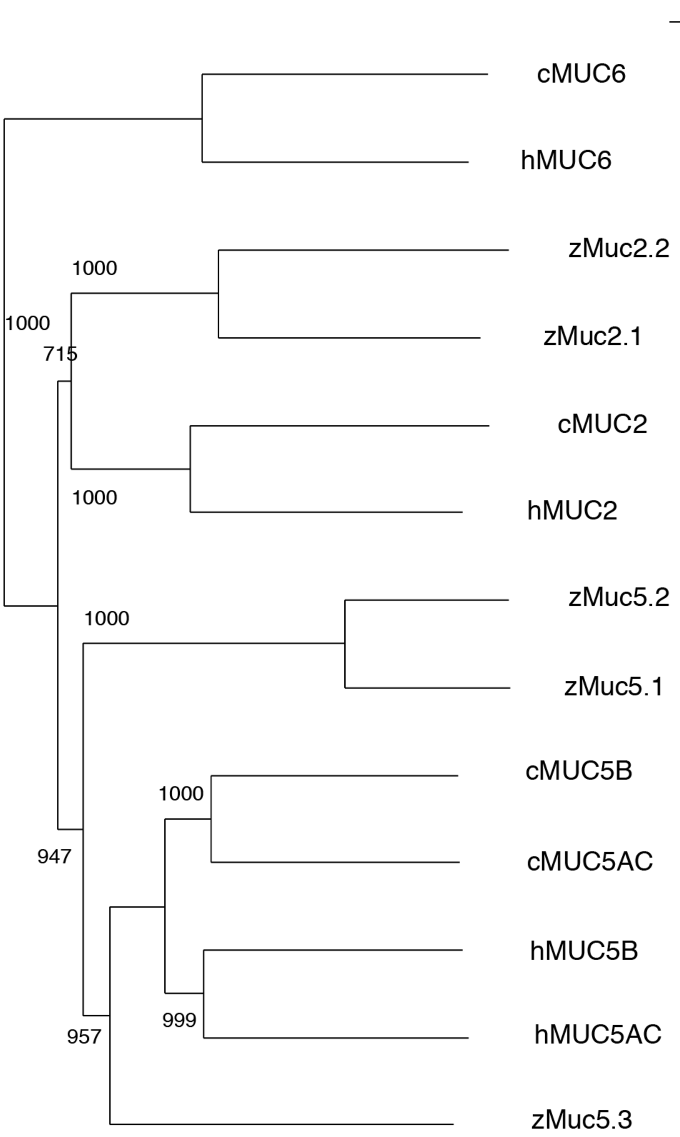
Supplementary Addendum S1. Protein sequences of secreted gel-forming mucins in zebrafish used for the construction of the protein domain models.

Supplementary Addendum S2. Protein sequences used for the construction of the phylogenetic trees.

Supplementary Addendum S3. Sequences of the verification transcripts.

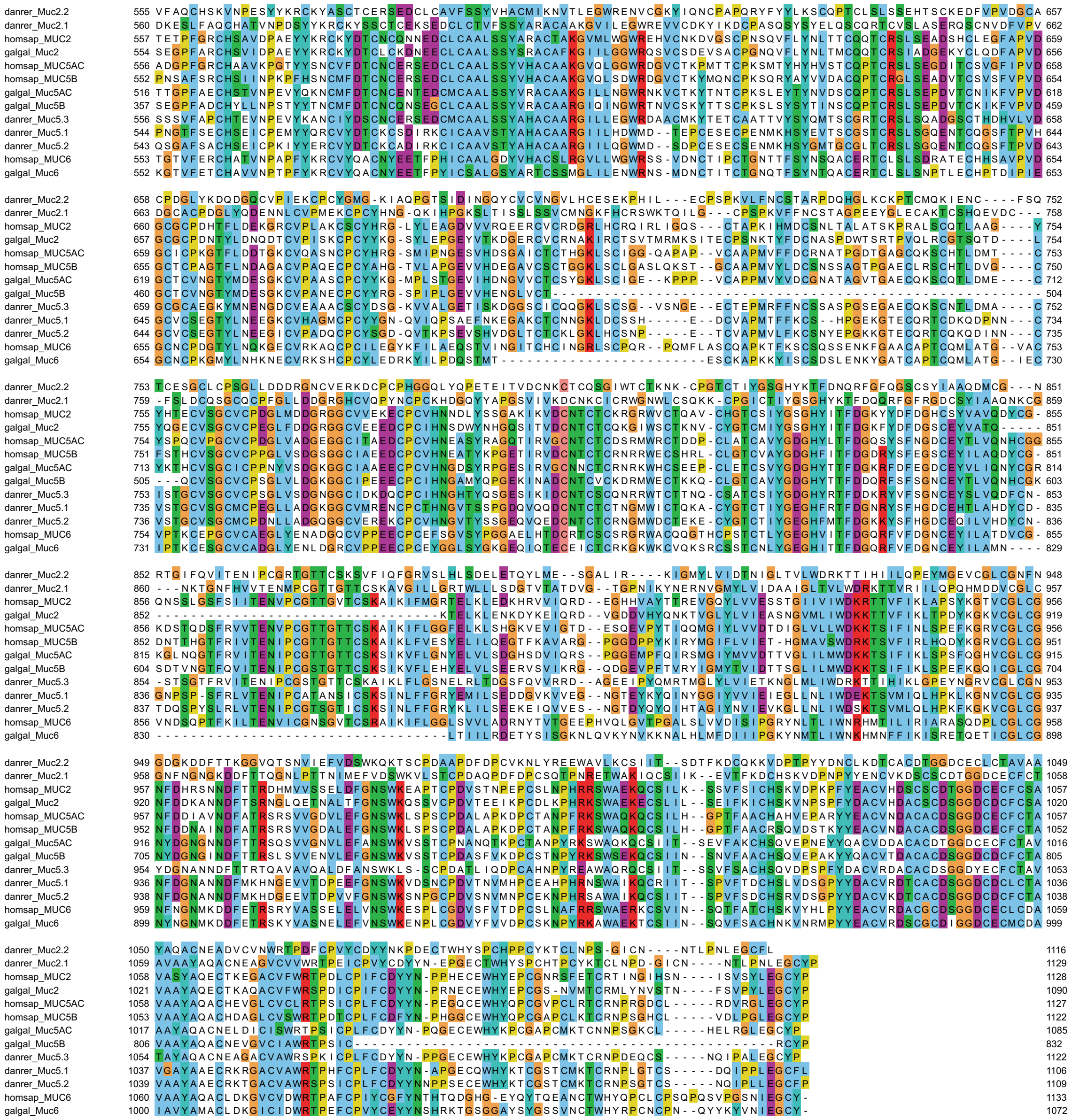


Supplementary Figure S1. Genomic structure of five predicted polymeric secreted mucins in zebrafish. Four of the mucins have been assigned across chromosomes 25 and 7. For *muc2.1*, the genomic position is unknown and hence, it is arbitrarily depicted left to the mucin gene cluster on chr25, relative to the position of the gap in the current genome assembly. The top bars show the genomic locations of computationally predicted VWD and PTS coding domains. PTS domains (blue boxes) were predicted by PTSPRED [22, 23] and VWD domains (grey boxes) were predicted by Genewise [37]. Ensembl transcript models [36] (in red) for the corresponding genomic regions are aligned below. To confirm the computationally predicted mucin domains we experimentally verified selected regions of each mucin transcript with RT-PCR and sequencing (pointers above transcript models represent the positions of the amplification primers). In pink is an addition to the *muc2.2* transcript that emerged from our experiments and that was missing in the Ensembl transcript model.

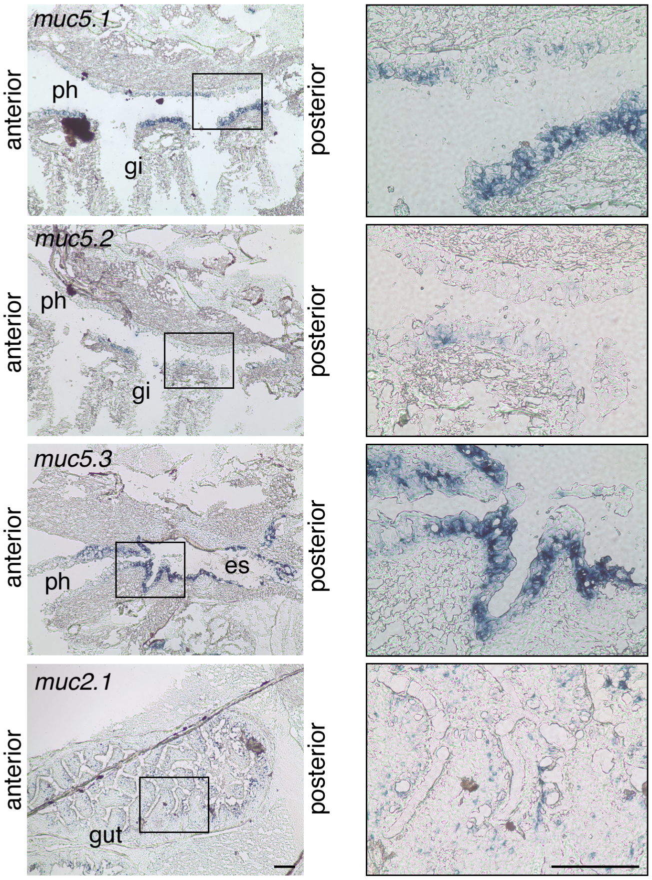


Supplementary Figure S2. Neighbor-joining phylogenetic tree of mucins in zebrafish, human, and chicken. The tree uses the same the N-terminal mucin sequences as the MrBayes tree in Fig. 1b. Numbers at branches represent bootstrap values. The tree reveals the same relationship among the mucins as MrBayes and is therefore a confirmation of their phylogenetic status. The only difference is seen with Muc5.3, which appears more distant from Muc5.1 and Muc5.2 and groups closer to the chicken and human mucins.

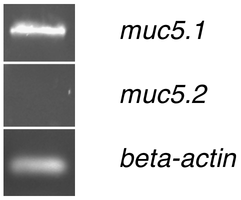
```
danrer_Muc2.2    1 STWESFHFKTFDGDIYQYPGLCEYNLVSDCTS--KQFSVDVKRTQ--TSSGPKISRIVVSINGKMFSISENNVMIDDQ-ITKLPVYESGILIE---ENTIYTN 95
danrer_Muc2.1    1 SMWGNFHFKTFDGDVYQFPGMCEYNLVSDCQSLIRQFSIYVKRTE--RSTGPKISRVSITINDIAIELTENQVNVNEA-KVTLPVHVSGILVE---ENTIYTR 97
homsap_MUC2      1 STWGNFHYKTFDGDVFRFPGLCDYNFASDCRGSYKEFAVHLKRGPGQAEAPAGVESILLTIKDDTIYLTRHLAVLNGA-VVSTPHYSPGLLIE---KSDAYTK 99
galgal_Muc2      1 STWGNNHFKTFDGDIYQFPGMCEYNFVSDCQDSYREFSVHIQRALNSNNHPE-IQYILITVTDFTVYLKPKLAVVDGR-IVKTPYYSSGLLIE---SNDIYTK 98
homsap_MUC5AC    1 STWGSFHYKTFDGDVFRFPGLCNYVFSEHCGAAYEDFNIQLRRSQ--ESAAPTLSRVLMKVDGVVIQLTKGSVLVNGH-PVLLPFSQSGVLIQ---QSSSYTK 97
homsap_MUC5B     1 STWGDFHYKTFDGDVFRFPGLCNYVFSEHCRAAYEDFNVQLRRGL--VGSRPVVTRVVIKAQGLVLEASNGSVLINGQ-REELPYSRTGLLVE---QSGDYIK 97
galgal_Muc5AC    1 STWGNFHFKNFDGDIFYFPGVCNYIFASNCKSPYEDFNIQIRRTM--VENATVITHVIMRLEGAVVELTRTSVLVNGK-QIQMPYSHMGVLIE---RSNSYIK 97
galgal_Muc5B     1 CTWGNFHFKTFDGDIFRYPGLCNYVFASHCNAPYEEFNIQIRRTL--VENILTIHRIAIKLEGVAIELTKDVVLINSN-RVQLPYSQAGITIE---RSSIYVK 97
danrer_Muc5.3    1 STWGNYHYKTFDGDYFQLPSTCNYILTSHCSSSYEDFNVQLRHEL--LNNEPTVSKITMKLQGTIIELSRDSLSVNGQ-TVTLPYSSAEVFIE---RVTSYIS 97
danrer_Muc5.1    1 STWGNFHFKTFDGHFFQLPDTCNYVLAVMCDAASSDFNIQMQRET--VNGSISFSTVLIKLDGTVIKVTDSDITMGEE-TVTVPTYKNGIKIE---GSPTSFK 97
danrer_Muc5.2    1 STWGNFHFKTFDGHFFQLPDTCNYVLAVMCDAAISDFNIQMQRDT--VNGSISFGSVIVKLDGTVIKITDSGIMMDDQ-VVSVPINQKGIKIE---GSPTSIK 97
homsap_MUC6      1 STWGAGHFSTFDHHVYDFSGTCNYIFAATCKDAFPTFSVQLRRGP-----DGSISRIIVELGASVVTVSEAIISVKDIGVISLPYTSNGLQITPFGQSVRLVA 98
galgal_Muc6      1 STWGGGHFSTFDKYQYDFTGTCNYIFATVCDESSPDFNIQFRRGL-----DKKIARIIIELGPSVIIVEKDSISVRSVGVIKLPYASNGIQIAPYGRSVRLVA 98

danrer_Muc2.2   96 IRSKTGLTVMWSKDGAVTVELDNKYQNKTCGLCGNFNRIRDEF-ITGNTEKIGYIEFGNKHKVFNPMEFCENPSETDEKNNWKAICGS--------------- 182
danrer_Muc2.1   98 LYSKMGITVMWNKDDAVMVELDSKYSNRTCGLCGDFNGIPVYNEFIQSGRTVGYTEFGNMHRVPNPTHQCEDPFENVDEQNVVDQCEK--------------- 185
homsap_MUC2    100 VYSRAGLTLMWNREDALMLELDTKFRNHTCGLCGDYNGLQSYSEFLSDGVLFSPLEFGNMQKINQPDVVCEDPEEEVAPASCSEHR----------------- 185
galgal_Muc2     99 VYAKLGLILIWNQEDALMVELDSKFGNQTCGLCGDYNGVPIYNEFISGVASYNSITYGNLQKISKPNDECEDPDETRALPSCNEHVSKPNCELRFPGCMETSL 201
homsap_MUC5AC   98 VEARLGLVLMWNHDDSLLLELDTKYANKTCGLCGDFNGMPVVSELLSHNTKLTPMEFGNLQKMDDPTEQCQDPVPEPPRNCSTGF------------------ 182
homsap_MUC5B    98 VSIRLVLTFLWNGEDSALLELDPKYANQTCGLCGDFNGLPAFNEFYAHNARLTPLQFGNLQKLDGPTEQCPDPLPLPAGNCTDEE------------------ 182
galgal_Muc5AC   98 VSAKLGLTFLWNEEDALLVELDKKYANQTCGLCGDFNGIPTNNEFISQNAKLTAVQFGNRQKIDGPTEQCDDPIPSAVLMNCSSEF----------------- 183
galgal_Muc5B    98 VVSKTDIVFMWNQRDSILLELHEKYANQTCGLCGDFNGIPVYNEFYSNNIRMTALQFGNMQKMDGPTEYCEDPTPSSLDN----------------------- 177
danrer_Muc5.3   98 IKAALGLVAMWNEDDSFMIRLDEKYRNQTCGLCGDFNGIPVYNEFVSGGSSLSVSDYGNLWKLDGPTETCTEDTMFLNENCGDEN------------------ 182
danrer_Muc5.1   98 ISNKHGVTVFWEEDNSLSIELPEKYQGQTCGLCGDFNGN-LADDITDNGP--------ATWKVSISTEICEEVTLPSTGPCDELSEQ---------------- 175
danrer_Muc5.2   98 IS-RYGMTIFWEEDNSIAIELAEKYKGLTCGLCGNYNGD-KDDDMPESGP--------ATWKISTPTESCEDVILPPKDQCDQN------------------- 171
homsap_MUC6     99 KQLELELEVVWGPDSHLMVLVERKYMGQMCGLCGNFDGKVTNEFVSEEGKFLEPHKFAALQKLDDPGEICTFQDIPSTHVRQAQH------------------ 183
galgal_Muc6     99 KLMEMELVVMWNNEDYLMVLTEKKYMGKTCGMCGNYDGYELNDFVSE-GKLLDTYKFAALQKMDDPSEICLSEEISIPAIPHKKY------------------ 182

danrer_Muc2.2  183 ----YQADCEELFRAKAWSSCAMSYHP--YVDACVYDRCFSK-----SSSLCATLSEYSRQCSRNGGTPSTWRTKDFCAVTCPYNTVHFESGSPCIDTCSSEN 274
danrer_Muc2.1  186 ----YRADCADLLEDEKWSSCSWVLDPEAYIKACTNDLCNRQPEDEDTTALCATLTEYSRQCSHAGGNPPAWRTAKFCNV-QCPYNMVHSESGSPCMDTCSHK 283
homsap_MUC2    186 ------AECERLLTAEAFADCQDLVPLEPYLRACQQDRCRCPGGD-T--CVCSTVAEFSRQCSHAGGRPGNWRTATLCPK-TCPGNLVYLESGSPCMDTCSHL 278
galgal_Muc2    202 HCEMLREECEKLLTSSAFADCHSRLNLEMYIQACMQDKCACKGNE-DSFCLCSTISEYSRQCSHAGGRPG--------------------------------- 270
homsap_MUC5AC  183 ------GICEELLHGQLFSGCVALVDVGSYLEACRQDLCFCEDTD-LLSCVCHTLAEYSRQCTHAGGLPQDWRGPDFCPQ-KCPNNMQYHECRSPCADTCSNQ 277
homsap_MUC5B   183 ------GICHRTLLGPAFAECHALVDSTAYLAACAQDLCRCP------TCPCATFVEYSRQCAHAGGQPRNWRCPELCPR-TCPLNMQHQECGSPCTDTCSNP 272
galgal_Muc5AC  184 -----ASICKTVLTSEAFTSCNALVDVQDYIDTCIQDLCHCDTSM-ADFCMCNTFAEYSRQCAHAGGQPLNWRTSELC------------------------- 255
galgal_Muc5B       -------------------------------------------------------------------------------------------------------
danrer_Muc5.3  183 -------FCRDIFSSPVFSDCRGLVSIDSFVQVCIQDLCHCNSSS-GAFCLCDTIAEYSRQCVHAGGKPEEWRTEYFCPH-PCPETMVYSQCGNPCTDTCTNS 276
danrer_Muc5.1  176 -----ASFCEEYLISPGFSGCYDVMDMRIFQKACVSDLCQCYGN---HDCLCNTLTEISRQCTHAGGQPGTWRTEQLCPK-TCPLNLQYLECGGPCRNTCSAP 269
danrer_Muc5.2  172 -----TMVCQQYLSSPGFSGCYDVMDMKIFEKACVSDMCQCYGS---HDCLCNTLTEISRQCTHAGGQPGTWRTEQLCPK-MCPINLQYMECGGPCKSTCSDP 265
homsap_MUC6    184 -----ARICTQLLT---LVAPECSVSKEPFVLSCQADVAAAPQPG-PQNSSCATLSEYSRQCSMVGQPVRRWRSPGLCSVGQCPANQVYQECGSACVKTCSNP 277
galgal_Muc6    183 -----AVICSQLLN---LVSPTCSVPKDGFVTRCQLDMQDCSEPG-QKNCTCSTLSEYSRQCAMSHQVVFNWRTENFCSVGKCSANQIYEECGSPCIKTCSNP 276

danrer_Muc2.2  275 TNTQCEEHNIDGCFCPPGTVFDDILNMGCIPREQCHCKHDRIYKPQEELHEVGQDCVCQNGKWNCKIKPVLGRCSVEEGSHFTTFDGQEFTFHGDCTYVLSKD 377
danrer_Muc2.1  284 DTNALCEEHNIDGCFCPPG-TVFDDISDT-GCIPAEKCQCKHDR-VYSSGEVLRKSEEECYCQEGSWVCMSIPGPGLCAVEEGSHFTTFDGKEFTFHGDCNYL 383
homsap_MUC2    279 EVSSLCEEHRMDGCFCPEG-TVYDDIGDS-GCVPVSQCHCRLHGHLYTPGQEITNDCEQCVCNAGRWVCKDLPCPGTCALEGGSHITTFDGKTYTFHGDCYYV 379
galgal_Muc2    271 --------------------TVYDDITEN-GCIPVSQCSCRLHGKGYSPGETITNECEECTCDSGRWTCKDLPCPGTCSVEGGSHIKTFDGKKYTFHGDCYYV 352
homsap_MUC5AC  278 EHSRACEDHCVAGCFCPEG-TVLDDIGQT-GCVPVSKCACVYNGAAYAPGATYSTDCTNCTCSGGRWSCQEVPCPDTCSVLGGAHFSTFDGKQYTVHGDCSYV 378
homsap_MUC5B   273 QRAQLCEDHCVDGCFCPPGSTVLDDITHS-GCLPLGQCPCTHGGRTYSPGTSFNTTCSSCTCSGGLWQCQDLPCPGTCSVQGGAHISTYDEKLYDLHGDCSYV 374
galgal_Muc5AC  256 -------------------RMVFDDINGA-GCIPRKECHCTYEGETYAPGTSFSSKCRSCTCAGGEWTCVTQSCPATCSIEGGSHITTYDEKYYSFYGDCSYV 338
galgal_Muc5B   178 ---------------------------------------CTYN-------------------------------------------------FH--------- 183
danrer_Muc5.3  277 ERGQVCAQHCIDGCFCPPG-TVYDDIHQS-GCIPYTECSCVLGGKIYASGENYTSNCRDCTCDQGQWDCYIKDCPHSCSVEGGSHITTFDGKAYTFHGDCTYV 377
danrer_Muc5.1  270 TANLMCTDHCVDGCFCPEG-TVEDDIGQS-GCVPVNECPCVHDGTVYKSGESYQQACKSCTCDAGHWTCTYLDCPGTCSVVGGSHVTTFDGKSFTFSGNCDYI 370
danrer_Muc5.2  266 TAHLMCKDHCVDGCFCPEG-TVEDDIGQG-GCVPVNECPCVHDGTVYKSGESYQQACKKCFCAAGHWTCTYLDCPGTCSVVGGSHVTTFDGKSFTFSGNCDYI 366
homsap_MUC6    278 QHS--CSSSCTFGCFCPEG-TVLNDLSNNHTCVPVTQCPCVLHGAMYAPGEVTIAACQTCRCTLGRWVCTERPCPGHCSLEGGSFVTTFDARPYRFHGTCTYI 377
galgal_Muc6    277 EYS--CSSHCTYGCFCPEG-TVLDDISKNRTCVHLEQCPCTLNGETYAPGDTMKAACRTCKCTMGQWNCKELPCPGRCSLEGGSFVTTFDSRSYRFHGVCTYI 376

danrer_Muc2.2  378 SVGSK-FIILGQISPCASLIYDTCLKSIDVIFNSDTVNRLVISEDGTVKHNAAVLLPYTPASRQFVVFSPSSFYILIDT-NFGLRVLVQLVPLMQVYITMDKR 478
danrer_Muc2.1  384 LSKDCEESK-FIILGQIVPCFTHETDTCLKSIVVLFDNDKKNPLFIKADGTVQHNAEVSLPYMTAD--FTVFRPSSFHIILQT-VF-GLQVQVQLVPLMQVYI 481
homsap_MUC2    380 LAKGDHNDS-YALLGELAPCGSTDKQTCLKTVVLLADKKKNAVVFKS-DGSVLLNELQV-NLPHVTASFSVFRPSSYHIMVSM-AIGVRLQVQLAPVMQLFVT 478
galgal_Muc2    353 LAKSAVNDT-HALLAELAPCGSSDGQTCLKTVVLLVDDKKNVVVFRS-DGSVSLNEMTV-NVPHVSASFSVFKPSSYYLIVHT-SFGLQLQIQLSPVMQLFVT 451
homsap_MUC5AC  379 LTKPCDSSA-FTVLAELRRCGLTDSETCLKSVTLSLDGAQTVVVIKA-SGEVFLNQIYT-QLPISAANVTIFRPSTFFIIAQT-SLGLQLNLQPVPTMQLFMQ 477
homsap_MUC5B   375 LSKKCADSS-FTVLAELRKCGLTDNENCLKAVTLSLDGGDTAIRVQA-DGGVFLNSIYT-QLPLSAANITLFTPSSFFIVVQT-GLGLQLLVQLVPLMQVFVR 473
galgal_Muc5AC  339 LTKLCDSND-FTVLGDIHKCGLTDTETCLKGVAISLSGGQTNIVIQP-SGSVFVNMIYT-QLPFSAANVTIFRPSSFFIILQT-TFGLQLQIQIVPLMQLFIN 437
galgal_Muc5B   184 ----CEDEI-FTILVELRKCGMTNTETCLKTVTLSMNKGQTLIEVKT-DGGVFMNSIYT-QLPISAANVTIFRPSSFFVIMHT-NIGIQVEIQITPIMQVFVR 278
danrer_Muc5.3  378 LSKHCSGTE-FTVLGDLVKCGLTETETCLKAITVALFDGITVINVDS-SGNVEVNRIVA-QLPLFTSSVSVFKPSSFYIIIEAPLIGLRLEIQLVPIMQLYIT 477
danrer_Muc5.1  371 LTKHSADSD-FAVVGNLAKCEASRTDTCLHSVTLVISG--TTISFSS-SGDVTLNNINSKELPFSIGPVSIFQPSSSFIIVDL--TSVRLEIQLAPVMQLYIV 467
danrer_Muc5.2  367 LTKHSNDSD-FAVVGNLAKCEPARTDTCLNSVTLVISG--TTIGFTS-DGVVTLNGNSPFNLPAVIGPVSIFQPSLSFIIADL--NSLRLEIQLAPVMQLYIV 463
homsap_MUC6    378 LLQSPQLPEDGALMAVYDKSGVSHSETSLVAVVYLSRQDKIVISQDE----VVTNNGEAKWLPYKTRNITVFRQTSTHLQMAT-SFGLELVVQLRPIFQAYVT 475
galgal_Muc6    377 LMKSSSLPHNGTLMAIYEKSGYSHSETSLSAIIYLSTKDKIVISQNE----LLTDDDELKRLPYKSGDITIFKQSSMFIQMHT-EFGLELVVQTSPVFQAYVK 474

danrer_Muc2.2  479 FQGK------------------------THGLCGNFNKVLADDMKTPQGLVEGTAVSFANAWEVE-TNC-S-RSDKSVIPCSHNSVSENYAEHWCYKIMDKSD 554
danrer_Muc2.1  482 TVDQSFQGK------------------------TCGLCGNFNKVLADDLKTPQGVVEGTAVSFANAWKAQ-SNCPDRSERMDDPCSYSSDSEHFAEHWCSKMK 559
homsap_MUC2    479 LDQASQGQ------------------------VQGLCGNFNGLEGDDFKTASGLVEATGAGFANTWKAQ-STCHDKLDWLDDPCSLNIESANYAEHWCSLLKK 556
galgal_Muc2    452 ADQSVQDIQGKCPLKQEAKPCISTAVKSYILYMGGLCGNFNGMEGDDFKTTNGLVEATGSAFANTWKAQ-PTCADQAEKLEDPCTLSIESANYAEHWCSLLKN 553
homsap_MUC5AC  478 LAPKLRGQ------------------------TCGLCGNFNSIQADDFRTLSGVVEATAAAFFNTFKTQ-AACPNIRNSFEDPCSLSVENEKYAQHWCSQLTD 555
homsap_MUC5B   474 LDPAHQGQ------------------------MCGLCGNFNQNQADDFTALSGVVEATGAAFANTWKAQ-AACANARNSFEDPCSLSVENENYARHWCSRLTD 551
galgal_Muc5AC  438 LDPSQKGQ------------------------TCGLCGNFNDIQTDDFKTTSGVIEGTSAAFGNTWKTR-ADCPDAKITFENPCTLSIENDKYAQHWCGLLSD 515
galgal_Muc5B   279 LDPVFKDQ------------------------TCGLCGNFNNVQTDDFKAISGVVEGTAAAFANTWKTQ-ASCPDIQSSFENPCALSIDNEKYAKHWCGLLTN 356
danrer_Muc5.3  478 ANSVYQGE------------------------TCGLCGNYNNMEADDFTTIGGLREGTALDFANTWKTR-ASCPDVKRSFENPCSLSTENEQYAQHWCSLLSD 555
danrer_Muc5.1  468 ASTEEKGK------------------------MTGLCGNYNDVQTDDFKTESGIIEGTPTSFVNFWKQN---CPDLEITFDNPCSLNMDTEKLAKDWCSRLTN 543
danrer_Muc5.2  464 ASTEEKGK------------------------MTGLCGNYNDVQSDDFKTDLGITEGTAISFANFWKKTPYSCPDLENTFDNPCSLSVDTEKLAKDWCSRLTN 542
homsap_MUC6    476 VGPQFRGQ------------------------TRGLCGNFNGDTTDDFTTSMGIAEGTASLFVDSWRAG--NCPAALERETDPCSMSQLNKVCAETHCSMLLR 552
galgal_Muc6    475 VSAQFQGR------------------------TLGLCGNYNGDTTDDFMTSMDITEGTASLFVDSWRAG--NCLPAMERETDPCALSQLNKISAETHCSILTK 551
```



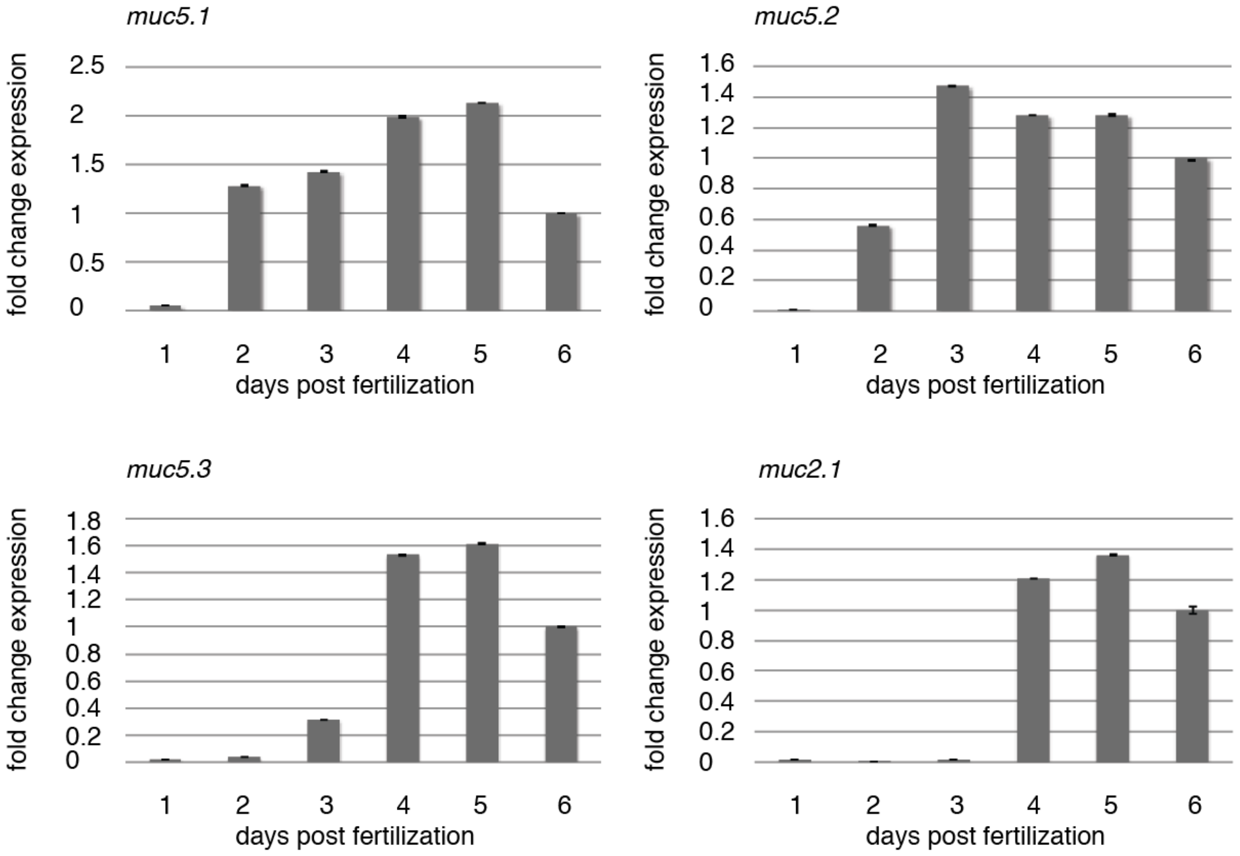
Supplementary Figure S3. Alignment of gel-forming secreted mucins from zebrafish, human, and chicken. Mucin sequences containing the three N-terminal VWD domains were aligned using ClustalW and visualized in Jalview. These sequences are the same as those used for the construction of the phylogenetic trees in Fig. 1b and Supplementary Fig. S2.



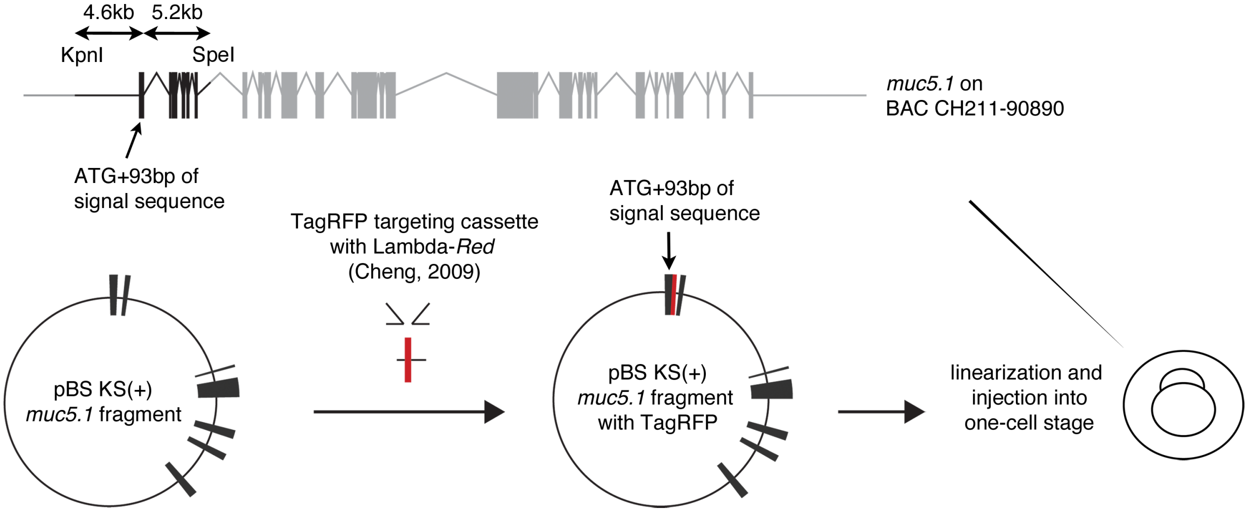
Supplementary Figure S4. *In situ* hybridization of the secreted mucins in sagittal sections of adult zebrafish. Image on the right shows an enlargement of the boxed area on the left. ph - pharynx, gi - gills, es - esophagus. DIG-UTP antisense RNA probes were used and positives are visualized in blue via NBT/BCIP reaction. The *in situ* hybridization signal suggests the localization of *muc5.1* and *muc5.2* in the pharynx, *muc5.3* in the esophagus, and *muc2.1* in the gut, supporting the RT-PCR data on isolated tissues in Fig 1c. Scale bar: 100 μm.



Supplementary Figure S5. Differential expression of *muc5.1 and muc5.2* in gill lamellae. The gill lamellae were separated from the gill arches. The total RNA was extracted from the lamellae, transcribed into cDNA and mucin transcripts were amplified by PCR. *muc5.1* but not *muc5.2* is expressed in the lamellar part of the zebrafish gills.



Supplementary Figure S6. qRT-PCR analysis of mucin gene expression during zebrafish development. The expression of individual mucin genes was normalized to *beta-actin* expression and compared to the expression at day six for each individual mucin. The timing of expression of each mucin correlates with the development of the respective organs in which they are found in the adult fish, namely the skin for *muc5.1* and *muc5.2*, the esophagus for *muc5.3* and the gut for *muc2.1* [27, 28].



Supplementary Figure S7. Generation of the fluorescent mucin reporter *muc5.1:S-RFP*. The reporter consists of the Tag-RFP (Tag-labeled Red Fluorescent Protein) fused in frame with the mucin signaling sequence downstream of the *muc5.1* promoter (S). A 9.8kb KpnI/SpeI fragment, which contains the mucin promoter region and part of the 5' region of *muc5.*1, was excised from BAC CH211-90890 and cloned into the pBSII KS(+) vector. The Tag-RFP targeting cassette was generated with high fidelity PCR and inserted behind the mucin signaling sequence in the first exon by Lambda-*Red* homologous recombination (see [42, 43] and Methods for details). The resulting construct contains the *muc5.1* promoter region (4.6 kb in the promoter and 5.2 kb in the 5' genomic sequence) and RFP fused in-frame with ATG. RFP is terminated with the stop codon.

Supplementary Table S1. Additional mucin-like fragments from the zebrafish genome identified in the bioinformatics searches. Their classification was determined with MrBayes. It appears that the zebrafish genome might contain at least one more *muc5*- and one more *muc2*-like mucin.

Ensembl ID	**Position**	**Genomic size (bp)**	**Domains**	**Mucin family**
ENSDARG00000089256	Zv9_NA551	13024	Cys-knot	*muc5*
ENSDARG00000090240	Zv9_NA551	2222	VWD	*muc5*
ENSDARG00000060302	Zv9_NA551	500	VWD	*muc5*
ENSDARG00000089023	Zv9_NA447	9544	VWD	*unknown*
ENSDARG00000076684	Zv9_NA447	18758	VWD	*muc2*
ENSDARG00000073938	Zv9_NA774	22126	PTS, VWD, Cys-knot	*muc2*
ENSDARG00000089990	chr25:8,583,497-8,592,403	8907	none	*muc5*

Supplementary Addendum S1. Protein sequences of secreted gel-forming mucins in zebrafish used for the construction of the protein domain models.

```
>danrer_Muc5.1 | chr25 cluster | IDENTICAL TO ENSEMBL : ENSDARG00000070331 |
ENSDART00000153708| ENSDARP00000127611

MGFDSEGTVRMPQMWMLRWVFLLAGLQSIQAGFMGDYRDMENPMTPMWPTTPAISMTSVPVITVEPNPDHRSTICST
WGNFHFKTFDGHFFQLPDTCNYVLAVMCDAASSDFNIQMQRETVNGSISFSTVLIKLDGTVIKVTDSDITMGEETVT
VPTYKNGIKIEGSPTSFKISNKHGVTVFWEEDNSLSIELPEKYQGQTCGLCGDFNGNLADDITDNGPATWKVSISTE
ICEEVTLPSTGPCDELSEQASFCEEYLISPGFSGCYDVMDMRIFQKACVSDLCQCYGNHDCLCNTLTEISRQCTHAG
GQPGTWRTEQLCPKTCPLNLQYLECGGPCRNTCSAPTANLMCTDHCVDGCFCPEGTVEDDIGQSGCVPVNECPCVHD
GTVYKSGESYQQACKSCTCDAGHWTCTYLDCPGTCSVVGGSHVTTFDGKSFTFSGNCDYILTKHSADSDFAVVGNLA
KCEASRTDTCLHSVTLVISGTTISFSSSGDVTLNNINSKELPFSIGPVSIFQPSSSFIIVDLTSVRLEIQLAPVMQL
YIVASTEEKGKMTGLCGNYNDVQTDDFKTESGIIEGTPTSFVNFWKQNCPDLEITFDNPCSLNMDTEKLAKDWCSRL
TNPNGTFSECHSEICPEMYYQRCVYDTCKCSDIRKCICAAVSTYAHACAARGIILHDWMDTEPCESECPENMKHSYE
VTSCGSTCRSLSGQENTCQGSFTPVHGCVCSEGTYLNEEGKCVHAGMCPCYYGNQVIQPSAEFNKEGAKCTCNNGKL
DCSSHEDCVAPMTFFKCSHPGEKGTECQRTCDKQDPNNCVSTGCVSGCMCPEGLLADGKGGCVMRENCPCTHNGVTS
SPGDQVQQDCNTCTCTNGMWICTQKACYGTCTIYGEGHFRTFDGNRYSFHGDCEHTLAHDYCDGNPSPSFRLVTENI
PCATANSICSKSINLFFGRYEMILSEDDGVKVVEGNGTEYKYQINYGGIYVVIEIEGLLNLIWDEKTSVMIQLHPKL
KGNVCGLCGNFDGNANNDFMKHNGEVVTDPEEFGNSWKVDSNCPDVTNVMHPCEAHPHRNSWAIKQCRIITSPVFTD
CHSLVDSGPYYDACVRDTCACDSGGDCDCLCTAVGAYAAECRKRGACVAWRTPHFCPLFCDYYNAPGECQWHYKTCG
STCMKTCRNPLGTCSDQIPPLEGCFLQCPSERPYLREETMKCVTEEGCKDCFYDGKVYPRGSVIYNTTDGSGACFSA
KCDYNGEIIREIKTDCIITTTPFTFSTPPSTTVTITSTETPSTITSTITPTTGTSTPTSTETPITTTSITSPTTETS
TSTSTETPPTPASTTSPTTETSTSTSTETPSTITSTVTPTTETSTPTSTETVPTPSTTTATTTPTSTTTPICVPRCE
WSDWIDSNKPNTEPEGFETESIPLLWNSRRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGL
PPICYNYEIRVYCCKENCPSDTTTSTPSTSTETPSTTTSTTSPTTETSTPTSTETPSTITSTVTPTTGTSTPTSTET
PPTPASTTSPTTETSTPTSTETPSTTASTTSPTTETSTPTSTETPSTITSTVTPTTETSTPTSTETVPTPSTTTATT
PTSTTTPICVPRCEWSDWIDSNKPNTEPEGFETESIPLLWNSKRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTS
FGLSCSNEDNAEGLPPICYNYEIRVYCCKENCPSDTTTSTPSTSTETPSTTSPTTETLTPTSTETPSTITSTVTPTT
GTLTPTSTETPSTTASTTSPTTETSTPTSTETPSTTASTTSPTTETSTPTSTETPSTTASTTSPTTETLTPTSTETP
STITSTVTPTTETSTPTSTETVPTPSTTTATTTPTSTTTPICVPRCEWSDWIDSNKPNTEPEGFETESIPLLWNSRR
ITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICYNYEIRVYCCKENCPSDTTPSTPSTS
TETPSTTTSTTSPTTETSTPTSTETPSTITFTITPTTGTSTPTSTETPPTPASTTSPTTETSTPTSTETPSTTASTT
SPTTETSTPTSTETPSTITSTVTPTTETSTPTSTETVPTPSTTTATTTPTSTTTPICVPRCEWSDWIDSNKPNTEPE
GFETESIPLLWNSRRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICYNYEIRVYCCK
ENCPSDTTPSTPSTSTETPSTTTSTTSPTTETSTPTSTETPSTITFTITPTTGTSTPTSTETPPTPASTTSPTTETS
TPTSTETPSTITSTVTATSETLTPTSTETVPTPSTTTATTATPTSTTTPICVPRCEWSDWIDSNKPNTEPEGFETES
IPLLWNTKRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICYNYEIRVYCCKDNCPSD
TTTSTPSTSTETPSTTASTTSPTTETSTPTSTETPSTITSTVTPTTGTSTPTSTETPITTTSITSPTTETSTPTSTE
TPSTTASTTSPTTETSTPTSTETPSTITSTVTPTTETLTPTSTETVPTPSTTTATTTPTSTTTPICVPRCEWSDWID
SNKPNTEPEGFETESIPLLWNSKRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICYN
YEIRVYCCKDNCPSDTTTSTPSTSTETPSTTASTTSPTTETSTPTSTETPSTITSTVTPTTGTLTPTSTETPSTITS
TVTPTTETSTPTSTETVPTPSTTTATTTPTSTTTPICVPRCEWSDWIDSNKPNTEPEGFETESIPLLWNSRRITCQS
PEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICYNYEIRVYCCKENCPSDTTTSTPSTSTETPS
TTTSTTSPTTETSTPTSTETPSTITSTVTPTTGTSTPTSTETPPTPASTTSPTTETSTSTSTETPPTPASTTSPTTE
TSTSTSTETPSTITSTVTPTTETSTPTSTETVPTPSTTTATTTPTSTTTPICVPRCEWSDWIDSNKPNTEPEGFETE
SIPLLWNSRRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICYNYEIRVYCCKENCPS
DTTTSTPSTSTETPSTTTSTTSPTTETSTPTSTETPSTITSTVTPTTGTSTPTSTETPPTPASTTSPTTETSTSTST
ETPPTPASTTSPTTETSTSTSTETPSTITSTVTPTTETSTPTSTETVPTPSTTTATTTPTSTTTPICVPRCEWSDWI
DSNKPNTEPEGFETESIPLLWNSRRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLSCSNEDNAEGLPPICY
NYEIRVYCCKENCPSDTTTSTPSTSTETPSTTTSTTSPTTETSTPTSTETPSTITSTVTPTTGTSTPTSTETPITTT
SITSPTTETSTPTSTETPSTTASTTSPTTETSTPTSTETPSTITSTVTPTTETLTPTSTETVPTPSTTTATTTPTST
TTPICVPRCEWSDWIDSNKPNTEPEGFETESIPLLWNSKRITCQSPEEVECRAVNYPQKTLEDLGQTVYCNTSFGLS
CSNEDNAEGLPPICYNYEIRVYCCKENCPSDTTTSTPSTSTETPSTTTSTTSPTTETSTPTSTETPSTITSTVTPTT
```

GTSTPTSTETPITTTSITSPTTETSTPTSTETPSTTASTTSPTTETATPTSTETPSTITSTVTPTTETSTPTSTETV
PTPSTTTATTTPTSTTTPICVPKCEWSDWIDSNKPNTEPEGFETESIPLLWNSKRITCQSPEEVECRAVNYPQKTLE
DLGQTVYCNTSFGLSCSNEDNAEGLPPICYNYEIRVYCCKYNCPSDTTTSTPYTSTETPSTTASTTSPTTETSTPTS
TETPSTITSTVTPTTGTLTPTSTETPSTTASTTSPTIETSTPTSTETPSTITSTVTPTTGTSTPTSTETPITTTSIT
SPTTETSTPTSTETPSTTASTTSPTTETPTPTSTETPSTITSTITPTTGTSTPTSTETPITTTSITSPTTETSTPTS
TETPSTTASTTSPTTETATPTSTETPSTITSTVTPTTGTLTPTSTETPSTTASTTSPTTETLTPTSTETPSTITSTV
TPTTGTSTPTSTETPITTTSITSPTTETSTPTSTETPSTTASTTSPTTETSTPTSTETPSTITSTITPTTGTSTPTS
TETPSTTASTTSPTTETSTPTSTETPSTTASTTSPTTEISTPTSTETPSLGTSTVTPTTGTSAPTSTETPSTTTSIT
SPTTETSTPTSTETPSTITSTVTPTTETSTPTSTETTSTTIPVTTIPTETPSTVVTVTTLSTETPSTTVSVTTIPTE
TPSTIVTVTTISTETPSTTVTVTPISTKSPSTIVTATPKSTKTPSITIGSTTTMSTEKPSTSSKSTTTSTSTKPVTT
TTKFTVHTGPVRTTPNPEQPSTQCNCTYLNAIFLPGSSIYNKTDKAGWCFSAHCNSSCSIIIQAEPCATTTPPTTTP
TTTPTTTTSKPNKNDCTQVQRKNMESWIENCKNYTCKNGKVTSVPLQCDKSDKVMPTCTNGIKAKKVSYNNGCCSRY
ECECKCSGWGDPHYQMFDGTYYDFQGNCTYVLFKEIVPRYNISAYVKNYLCDLQNNLACPEYVIVNYKSYSIKLTSN
TKEIEVYVNDDLKQLTYVNNNFFITTSGIGVILNITEINVEITVNHQGFVINLPFSYFSGNTEGQCGVCDNNRTNDC
RRPDGQIDKSCENMAQLWMVPPGCQPPKPIPKPTPPPVICNPSICEYIKGDMFKECHDHIPYESYYESCKFDVCRMG
NRSIGCTSLEAYAALCGKEGICVDWRTSDDLKQTCAHKCPSHKVYNACGPKVEKTCSTRYNDLFVEKDCQGESCHQT
YMEGCFCPDNTYLVSSTTDYCTAYCDCIGPDGLPREPGDTWTMDCYKYTCSNETFGITKEFVGCPTIVPCGPEQKLI
IDNCCPTCVCNLEVCIHKKCDVGFELAEHANEGSCCPPCVPKDVCVYNNTEYQPGSTIYPDPCVECNCDMEKDPETQ
LHTVSCVSKICEPCSEGFELVEQKGECCGTCKPNACIYVSDNITHILSEGESYNYKCESVTCQQVNGTYIIERTIPT
CPEINPDQCVPGTLGLDEDGCCNTCELKNCVRVKNMTDVTVNDCKSINPIEVTSCSGNCDTESMYSMGANTMMHSCS
CCKETKTSVKKVTLKCADGSEIPHDYVYIESCRCTPITCENQKTSG

>danrer_Muc5.2 | chr25 cluster | IDENTICAL TO ENSEMBL ENSDARP00000075869 ENSDART00000081428 ENSDARG00000058556 comment: missing parts of pts domain as compared to genewise result below

CSTWGNFHFKTFDGHFFQLPDTCNYVLAVMCDAAISDFNIQMQRDTVNGSISFGSVIVKLDGTVIKITDSGIMMDDQ
VVSVPINQKGIKIEGSPTSIKISRYGMTIFWEEDNSIAIELAEKYKGLTCGLCGNYNGDKDDDMPESGKCHVFVVSS
TWKISTPTESCEDVILPPKDQCDQNTMVCQQYLSSPGFSGCYDVMDMKIFEKACVSDMCQCYGSHDCLCNTLTEISR
QCTHAGGQPGTWRTEQLCPKMCPINLQYMECGGPCKSTCSDPTAHLMCKDHCVDGCFCPEGTVEDDIGQGGCVPVNE
CPCVHDGTVYKSGESYQQACKKCFCAAGHWTCTYLDCPGTCSVVGGSHVTTFDGKSFTFSGNCDYILTKHSNDSDFA
VVGNLAKCEPARTDTCLNSVTLVISGTTVSSRFCILTTLFIVPGPVSIFQPSLSFIIADLNSLRLEIQLAPVMQLYI
VASTEEKGKMTGLCGNYNDVQSDDFKTDLGITEGTAISFANFWKKTPYSCPDLENTFDNPCSLSVDTEKLAKDWCSR
LTNQSGAFSACHSEICPKIYYERCVYDTCKCADIRKCICAAVSTYAHACAARGIILQGWMDSDPCESECSENMKHSY
GMTGCGLTCRSLSGQENTCQGSFTPVDGCVCSEGTYLNEEGICVPADQCPCYSGDQVTKPSEVSHVDGLTCTCKLGK
LHCSNPESKFQPCVAPMVLFKCSNYEPGKKGTECQRTCQKQDINNCVSTGCVSGCMCPDNLLADGQGGCVEREKCPC
VHNGVTYSSGEQVQEDCNTCTCRNGMWDCTEKECYGTCTIYGEGHFMTFDGKKYSFHGDCEQILVHDYCNTDQSPYS
LRLVTENIPCGTSGTICSKSINLFFGRYKLILSEEKEIQVVESNGTDYQYQIHTAGIYNVIEVKGLLNLIWDSKTSV
MLQLHPKFKGKVCGLCGNFDGNANNDFMKHDGEEVTDPVVFGNSWKSNPGCPDVSNVMNPCEKNPHRSAWAIKQCSI
ITSPVFSDCHSRVDSGPYYDACVKDTCACDSGGDCDCFCTAVAAYAAECRKTGACVAWRSPSICPLFCDYYNNPPSE
CEWHYKTCGSTCMKTCRNPSGTCSNQIPLLEGCFPQCPSERPFLREDNRKCVTEAECLSLCTYDGKIYTTGDVMYDT
TDGNGTCFTAVCGSNGEIIRSINKSTTTSAPTTPTTPEVTTANTAETPLTETTTCYCTFANQTFPPGKNVIIKTDKA
GWCYFAYCNSNCVSEITQKPCPQTTTIPLSTVPNDCIDVNKKNGESWNEGCTTKNCINGKVTTSPVQCDSADSDIPT
CANGLKPEKVYYNNGCCTKYECSCQCSGWGDPHYSTFDGTYYVFQGICDYVLVQEIIPRYNISVHVKNNYCEATNKY
ACQDSVIVKYKGYEIKLASNTEQIQLNTLSVSVNDVVKVPTFLNEDFSITTSGMAVILNIKEIKSEIVVSHQGFKIN
LPFSYFKGNTEGQCGVCDNDATNDCRRPDGTIDQSCEQMAALWADSPGCESPTPHPSEPIPTCTAQVCEVIKSDLFK
SCHDIVPYQSYYEACKYDVCYKRNDTMACASLEAYAQLCGLKSVCVDWRGSDDLNGQCGYNCGDHKVYQACGPKVEK
TCSTRYNEMFADNGDQTFMEGCFCPENTYLLSSTTDQCTPTCDCIGPDGLPRQPGDSWVVNCTTYKCSSESFGVVQE
PMNCPKIKPCHEGYKSVIENCCPTCVCDTSLCPMKCDVGYELSEEIPEHHCCPHCVPKDVCVHNNTEYQPGVKIPTE
PCMECHCHMDKDKDMDEDRHSKLHTVTCVKKCSKGFELVEQEGECCGTCKQKSCIYTAPDNTTHTLQPGEVKSHHKC
ETLTCREIDGVFVTEKIKPKCPDFNPDDCEAGTERFDDDGCCKICKPKICIVRKNTTHVNANDCKSISPIEVTSCSG
HCGTQSMYSMENSIMMHSCSCCREEKFVTRQVKLKCADGREIVHDYVYIESCTCTPTKCGSIKT

>danrer_Muc5.2_alt_genewise | chr25 cluster | obtained by genewise alignment with muc5.3 protein to chr25 region encoding muc5.2.

```
VEPKPDHQSRICSTWGNFHFKTFDGHFFQLPDTCNYVLAVMCDAAISDFNIQMQRDTVNGSISFGSVIVKLDGTVIK
ITDSGIMMDDQVVSVPINQKGIKIEGSPTSIKISRYGMTIFWEEDNSIAIELAEKYKGLTCGLCGNYNGDKDDDMPE
SGPATWKISTPTESCEDVILPPKDQCDQNTMVCQQYLSSPGFSGCYDVMDMKIFEKACVSDMCQCYGSHDCLCNTLT
EISRQCTHAGGQPGTWRTEQLCPKMCPINLQYMECGGPCKSTCSDPTAHLMCKDHCVDGCFCPEGTVEDDIGQGGCV
PVNECPCVHDGTVYKSGESYQQACKKCFCAAGHWTCTYLDCPGTCSVVGGSHVTTFDGKSFTFSGNCDYILTKHSND
SDFAVVGNLAKCEPARTDTCLNSVTLVISGTTIGFTSDGVVTLNGNSPFNLPAVIGPVSIFQPSLSFIIADLNSLRL
EIQLAPVMQLYIVASTEEKGKMTGLCGNYNDVQSDDFKTDLGITEGTAISFANFWKKTPYSCPDLENTFDNPCSLSV
DTEKLAKDWCSRLTNQSGAFSACHSEICPKIYYERCVYDTCKCADIRKCICAAVSTYAHACAARGIILQGWMDSDPC
ESECSENMKHSYGMTGCGLTCRSLSGQENTCQGSFTPVDGCVCSEGTYLNEEGICVPADQCPCYSGDQVTKPSEVSH
VDGLTCTCKLGKLHCSNPETCVAPMVLFKCSNYEPGKKGTECQRTCQKQDINNCVSTGCVSGCMCPDNLLADGQGGC
VEREKCPCVHNGVTYSSGEQVQEDCNTCTCRNGMWDCTEKECYGTCTIYGEGHFMTFDGKKYSFHGDCEQILVHDYC
NTDQSPYSLRLVTENIPCGTSGTICSKSINLFFGRYKLILSEEKEIQVVESNGTDYQYQIHTAGIYNVIEVKGLLNL
IWDSKTSVMLQLHPKFKGKVCGLCGNFDGNANNDFMKHDGEEVTDPVVFGNSWKSNPGCPDVSNVMNPCEKNPHRSA
WAIKQCSIITSPVFSDCHSRVDSGPYYDACVKDTCACDSGGDCDCFCTAVAAYAAECRKTGACVAWRSPSICPLFCD
YYNNPPSECEWHYKTCGSTCMKTCRNPSGTCSNQIPLLEGCFPQCPSERPFLREDNRKCVTEAECLSLCTYDGKIYT
TGDVMYDTTDGNGTCFTAVCGSNGEIIRSINKCMTTTTPFTFTTPPQTPTPSSPETYTSVTAATPSTERPPTSELQR
PKTTTTTSGPTTKTSQGISTAETPSTVLPLPTSTATTTARTPPPTKTETPSTATTTAAPPTPTSPEVTTAVTAETPS
TGTPPSTTTETPSIATSTAAPPTPTAPEVTTAVTAETPSTGTPPSTKTETPSIATTTSAPPTPTAPEVTTAVTAETP
STEMPLHTTETPSTVTTTAASPTPTAPEVTTAVTAETPSTGTPPSTTTETPSIATSTAAQLTPTSPEVTTAVTAETP
STGTPPSTKTETPSTATTSAAPPTPPAPEVTTAVTAETSSTGTPPSTKTETPSTATTTAAPPTPPAPEVTTAVTAET
PSTGTPPSTKTETPSIVTTTAAPPTPTAPEVTTAVTAETPSTGTPPSTKTETPIIATTTAAPPTPTAPEVTTAVTAE
TPSTEMSLHTPTETPSTVTTTAAPPTPTAPEVTTAVTAETPSTGTPPSTKTETPSTATTTAAPPTPTAPEVTTAVTA
ETPSTGTPPSTKTETPSIATTTAAPPTPTAPEVTTAVTAKTPSTGTPPPTKTETPSTATTTAAPPTPTATEVTTAVT
AETPSTGTPPSTTTETPSIATSTAAPPTPTSPEVTTAVTAETPSTGTPPSTKTETPSTATTTAAPPTPTATEVTTAV
TAETPSTGTPPSTKTETPSTATTTAAPPTPPAPEVTTAVTAETPSTGTPPSTKTETPSIATTTAAPPTPTSPEVTTA
VTAETPSTGTPPSTTTETPSTVTTTAAPPTPTAPEVTTAVTPSTEMPLHTTEITETLSTRTPPSTKTETPSTATTTS
APTTPTTPEVTTANTAETPLTETTTCYCTFANQTFPPESVIYNKTDKAGWCYFAYCNSNCVSEITQKPCPQTTTIPL
STVPNDCIDVNKKNGESWNEGCTTKNCINGKVTTSPVQCDSADSDIPTCANGLKPEKVYYNNGCCTKYECSCQCSGW
GDPHYSTFDGTYYVFQGICDYVLVQEIIPRYNISVHVKNNYCEATNKYACQDSVIVKYKGYEIKLASNTEQIQVSVN
DVVKVPTFLNEDFSITTSGMAVILNIKEIKSEIVVSHQGFKINLPFSYFKGNTEGQCGVCDNDATNDCRRPDGTIDQ
SCEQMAALWADSPGCESPTPHPSEPIPTCTAQVCEVIKSDLFKSCHDIVPYQSYYEACKYDVCYKRNDTMACASLEA
YAQLCGLKSVCVDWRGSDDLNGQCGYNCGDHKVYQACGPKVEKTCSTRYNEMFADNGDQTFMEGCFCPENTYLLSST
TDQCTPTCDCIGPDGLPRQPGDSWVVNCTTYKCSSESFGVVQEPMNCPKIKPCHEGYKSVIENCCPTCVCDTSLCPM
KCDVGYELSEEIPEHHCCPHCVPKDVCVHNNTEYQPGVKIPTEPCMECHCHMDKDKD


>danrer_Muc2.2_chr25_cluster | chr25 cluster | IDENTICAL TO ENSEMBL
ENSDARG00000078994 |ENSDART00000112246| ENSDARP00000098665
comment: incomplete, without 4th vwd of muc2

ININGSFIPVQRDCLPPQNITCENHQPSVLLEDKHECCPYYVCDCFCQGWGDSHYITFDGLYYNYQGNCTYILMKEI
SPQFNLAIYIENVYCDPAENVSCPRSLIVSYNKQNITLRSHRLMGGANLEALEGNIMLALPYDNNGLRVISSDLNLL
LEIPKLNAYVLFGVSGFSINLPFQYFGKNTEGQCGTCNNNQNDDCPGSLANDCKIMAENWNTTVCTPPAPLPTPNSP
HVHPDCYLLNTTFEACHSHVPPENFFSACQYDSSNNPAEVCASLQSYASVCAMFGVCIYWRNYTSQCNIGCSEDKVF
LPCGSFEPPTCRDRQMESNITVSTEGCFCSQNTTLFTKDSGVCVPQPCGCLDSSGKPRRFNENFTHNCEDCVCDKAR
MSVICEPKKCPDVPPVNCTEPFMVVNVTDPSEPCCSRQVCNCSMCPFSEEKCEVGYELIVHPYGDFCPEIRCESKEV
CAYNNTEYKPGSKIHVGPCVKCNCEWDPWTELYQMNCSYEVCPDDGFQYMKKDGECCGSCVQVACIYDAPDKTTHVL
QVHLETFTNVPHISCMNVTCLKRNDVFTIKEEMKKCPSFNQDDCVEGTKQFDEDGCCPNCKICETRNCAPVKNVTRL
KVNNCTSTEDVEITSCAGHCDSKSTLDLHGYMMNSCSCCQAERSSIREVTLMCEDDWQKWLNIRHNYTYIETCTCTP
VISGLFQSTFIRS


>danrer_Muc2.2_chr25_cluster_genewise | chr25 cluster | Same as ENSEMBL
ENSDARG00000078994, but more complete version obtained by alignment of muc2.1
to genomic region of ENSDARG00000078994
```

SSMCSTWESFHFKTFDGDIYQYPGLCEYNLVSDCTSKQFSVDVKRTQTSSGPKISRIVVSINGKMFSISENNVMIDD
QITKLPVYESGILIEENTIYTNIRSKTGLTVMWSKDGAVTVELDNKYQNKTCGLCGNFNRIRDEFITGNTEKIGYIE
FGNKHKVFNPMEFCENPSETDEKNNWKAICGSYQADCEELFRAKAWSSCAMSYHPYVDACVYDRCFSKSSSLCATLS
EYSRQCSRNGGTPSTWRTKDFCAVTCPYNTVHFESGSPCIDTCSSENTNTQCEEHNIDGCFCPPGTVFDDILNMGCI
PREQCHCKHDRIYKPQEELHEVGQDCVCQNGKWNCKIKPVLGRCSVEEGSHFTTFDGQEFTFHGDCTYVLSKDSVGS
KFIILGQISPCASLIYDTCLKSIDVIFNSDTVNRLVISEDGTVKHNAAVLLPYTPASRQFVVFSPSSFYILIDTNFG
LRVLVQLVPLMQVYITMDKRFQGKTHGLCGNFNKVLADDMKTPQGLVEGTAVSFANAWEVETNCRSDKSVIPCSHNS
VSENYAEHWCYKIMDKSDVFAQCHSKVNPESYYKRCKYASCTCERSEDCLCAVFSSYVHACMIKNVTLEGWRENVCG
KYIQNCPAPQRYFYYLKSCQPTCLSLSSEHTSCKEDFVPVDGCACPDGLYKDQDGQCVPIEKCPCYGMGKIAQPGTS
IDINGQYCVCVNGVLHCESEKPHILECPSPKVLFNCSTARPDQHGLKCKPTCMQKIENCFSQTCESGCLCPSGLLDD
DRGNCVERKDCPCPHGGQLYQPETEITVDCNKCTCQSGIWTCTKNKCPGTCTIYGSGHYKTFDNQRFGFQGSCSYIA
AQDMCGNRTGIFQVITENIPCGRTGTTCSKSVFIQFGRVSLHLSDELETQYLMESGALIRKIGMYLVIDTNIGLTVL
WDRKTTIHIILQPEYMGEVCGLCGNFNGDGKDDFTTKGGVQTSNVIEFVDSWKQKTSCPDAAPDFDPCVKNLYREEW
AVLKCSIITSDTFKDCQKKVDPTPYYDNCLKDTCACDTGGDCECLCTAVAAYAQACNEADVCVNWRTPDFCPVYCDY
YNKPDECTWHYSPCHPPCYKTCLNPSGICNNTLPNLEGCFLTCPDDKPFYDDENQVCVKDCYISTTISTLTTTYTSP
LPTSSLMPSTTSEIPFTTSITPSTRSNSTFSTISVTTTPQNPSTTSETLSRTSTTPGTTTSIFTTTHSTTTPLTPST
T

>danrer_Muc2.1 | Zv9_NA774 | IDENTICAL TO ENSEMBL ENSDARG00000074142
|ENSDART00000046576 | ENSDARP00000046575 comment: no 4th vwd domain

VIFAVSPSNHVNSICSMWGNFHFKTFDGDVYQFPGMCEYNLVSDCQSLIRQFSIYVKRTERSTGPKISRVSITINDI
AIELTENQVNVNEAKVTLPVHVSGILVEENTIYTRLYSKMGITVMWNKDDAVMVELDSKYSNRTCGLCGDFNGIPVY
NEFIQSGRTVGYTEFGNMHRVPNPTHQCEDPFENVDEQNVVDQCEKYRADCADLLEDEKWSSCSWVLDPEAYIKACT
NDLCNRQPEDEDTTALCATLTEYSRQCSHAGGNPPAWRTAKFCNVQCPYNMVHSESGSPCMDTCSHKDTNALCEEHN
IDGCFCPPGTVFDDISDTGCIPAEKCQCKHDRVYSSGEVLRKSEEECYCQEGSWVCMSIPGPGLCAVEEGSHFTTFD
GKEFTFHGDCNYLLSKDCEESKFIILGQIVPCFTHETDTCLKSIVVLFDNDKKNPLFIKADGTVQHNAEVSLPYMTA
DFTVFRPSSFHIILQTVFGLQVQVQLVPLMQVYITVDQSFQGKTCGLCGNFNKVLADDLKTPQGVVEGTAVSFANAW
KAQSNCPDRSERMDDPCSYSSDSEHFAEHWCSKMKDKESLFAQCHATVNPDSYYKRCKYSSCTCEKSEDCLCTVFSS
YARACAAKGVILEGWREVVCDKYIKDCPASQSYSYELQSCQRTCVSLASERQSCNVDFVPVDGCACPDGLYQDENNL
CVPMEKCPCYHNGQKIHPGKSLTISSLSSVCMNGKFHCRSWKTQILGCPSPKVFFNCSTAGPEEYGLECAKTCSHQE
VDCFSLDCQSGCQCPFGLLDDGRGHCVQPYNCPCKHDGQYYAPGSVIVKDCNKCICRWGNWLCSQKKCPGICTIYGS
GHYKTFDQQRFGFRGDCSYIAAQNKCGNKTGNFHVVTENMPCGTTGTTCSKAVGILLGRTWLLLSDGTVTATDVGTG
PNIKYNERNVGMYLVIDAAIGLTVLWDRKTTVRIILQPQHMDDVCGLCGNFNGNGKDDFTTQGNLPTTNIMEFVDSW
KVLSTCPDAQPDFDPCSQTPNRETWAKIQCSIIKEVTFKDCHSKVDPNPYYENCVKDSCSCDTGGDCECFCTAVAAY
AQACNEAGVCVVWRTPEICPVYCDYYNEPGECTWHYSPCHTPCYKTCLNPDGICNNTLPNLEGCYPECPEDKPIFDE
KNQICVEECPTPPPPTTTPVPTTTTPVPTTTTIPPTTTPVPTTTTPVPPTTTPVPTTTTPVPPTTTPVPPTTTPVPT
TTTPVPPTIPPTTTPVPTTTPIPSTSPCIICEWSDWFNVYNPATGGWNDEETYKNIAASGGQICEQPEDIECRGVEA
KPGMSFEDYIHETKQVVQCNVSYGLVCKKDEQKFRPYKCFDYEIRVYCCYDCGSTTTEFTTSPTPIPTTTTQPPTTT
PEPTTTATVPPTTTPEPTTTTPIPTTTTEPPKPPTTTPTPTTTTELPTTTPEPTTTTPIPTTTT

>danrer_Muc5.3 | chr7 | IDENTICAL TO ENSEMBL ENSDARG00000089847
|ENSDART00000148421 | ENSDARP00000124449

XFTTMNLIPTATIEMVAGLTSQHNELICSTWGNYHYKTFDGDYFQLPSTCNYILTSHCSSSYEDFNVQLRHELLNNE
PTVSKITMKLQGTIIELSRDSLSVNGQTVTLPYSSAEVFIERVTSYISIKAALGLVAMWNEDDSFMIRLDEKYRNQT
CGLCGDFNGIPVYNEFVSGGSSLSVSDYGNLWKLDGPTETCTEDTMFLNENCGDENFCRDIFSSPVFSDCRGLVSID
SFVQVCIQDLCHCNSSSGAFCLCDTIAEYSRQCVHAGGKPEEWRTEYFCPHPCPETMVYSQCGNPCTDTCTNSERGQ
VCAQHCIDGCFCPPGTVYDDIHQSGCIPYTECSCVLGGKIYASGENYTSNCRDCTCDQGQWDCYIKDCPHSCSVEGG
SHITTFDGKAYTFHGDCTYVLSKHCSGTEFTVLGDLVKCGLTETETCLKAITVALFDGITVINVDSSGNVEVNRIVA
QLPLFTSSVSVFKPSSFYIIIEAPLIGLRLEIQLVPIMQLYITANSVYQGETCGLCGNYNNMEADDFTTIGGLREGT
ALDFANTWKTRASCPDVKRSFENPCSLSTENEQYAQHWCSLLSDSSSVFAPCHTEVNPEVYKANCIYDSCNCERSED
CMCAALSSYAHACAARGILLEGWRDAACMKYTETCAATTVYSYQMTSCGRTCRSLSQADGSCTHDHVLVDGCGCAEG
KYMNENGDCVEAAACSCYDSGKVVALGETISKDGGSCICQQGKLSCSGVSNGEECTEPMRFFNCSSASPGSEGAECQ
KSCNTLDMACISTGCVSGCVCPSGLVSDGNGGCIDKDQCPCIHNGHTYQSGESIKIDCNTCSCQNRRWTCTTNQCSA
TCSIYGDGHYRTFDDKRYVFSGNCEYSLVQDFCNSTSGTFRVITENIPCGSTGTTCSKAIKLFLGSNELRLTDGSFQ
VVRRDAGEEIPYQMRTMGLYLVIETKNGLMLIWDRKTTIHIKLGPEYNGRVCGLCGNYDGNANNDFTTRTQAVAVQA

```
LDFANSWKLSSCPDATLIQDPCAHNPYREAWAQRQCSIITSSVFSACHSQVDPSPFYDACVRDACACDSGGDYECFC
TAVTAYAQACNEAGACVAWRSPKICPLFCDYYNPPGECEWHYKPCGAPCMKTCRNPDEQCSNQIPALEGCYPQCPQE
QPVFDEDNMKCVKQEECGCFVDMEHYEVGEQVPTTENCQSCNCSSTGVICEFDINACRCQYNNDTYYYGDIIYHTTD
GLGGCITGICSVNGTIERSIDPCETTTVPTTTSSLITTPTPTTVFVFSTPESTTSEVSYTTKKPEQLTTGTTAEVIQ
TASTREIKHSTEQQVYVTTKPSTSLPTTMKKTSESETSTNILTTTMETSIAPITKTLPTVTDKPWTTLSTQSQSVST
RQTTEEQTTENELTTLETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQP
TTLTKTTAVPVTTLVSTTVTVESSTELEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPS
ILSTTSKSSTSTETTGVVTTTETVTENELSTSETTMSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSST
SVPLRKTSPAITSSQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLV
EGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTV
EELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQSPSSTTKPIKAET
TLSTTTAAIPSTTFSVVEGSTTVQPSAPSILSTTSKSSTATETTGVVTTTETVTENELSTSEKTLSTTRPPKVVTTG
QTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPPITSTQPTTVTKTTAVPVTTLVSTTVSVESSTKFEVSTQ
GPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTS
ETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTVTKTTAVPVTTLVS
TTVTVESSTELEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTSTETTG
VVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQP
TTVTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPS
ILSTTSKSSTVTETTGVVTSTETVTENELSTSETTLSTTRPPKVVTTGQTIEPTTTATTTVEELTPTSPESSTPSST
LVPVRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVTVESSTEFEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLV
EGSSTVQQPTETSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTV
EELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAET
TLSTTTAAIPSTTFSVVEGSTTVQPSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTG
QTIQPTTTATTTVEELTATSPESSTPSSTSLPVRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVTVESSTEFEVSTQ
GPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTS
ETTMSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTVTKTTAVPVTTLVS
TTVSVESSTEFEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETTG
VVTTTETVTENELSTSETTLSTTRPPQVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQP
TTVTKTTAVPVTTLVSTTVTVESSTELEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPS
ILSTTSKSSTSTETTGVVTTTETVTENELSTSETTMSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSST
SVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAAIPSTTFSV
VEGSTTVQPSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTSRPPKVVTTGQTIQPTTTATTTVEEL
MATSPESSTPSSTSLPVRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVTVESSTEFEVSTQGPSSTTKPIKAETTLS
TTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTSETTMSTTRPPKVVTTG
QTIQPTTTATTTVEELTATSPESSTPSSTSVPLKQTSPTITSTQPTTVTKTTAVPVTTLVSTTVSVESSTEFEVSTQ
SPSSTTKPIKAETTLSTTTAETPSTTVLLEGSSTVQQPTDTSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTS
ETTLSTTRPPKVVTTSQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVS
TTVSVESSTEFEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQKPTETSAPSILSTTSKSSTVTETTG
VVTTTETVTENELSTSETTLSTSRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSLPVRQTSPAITSTQP
TTVTKTTAVPVTTLVSTTVTVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPS
ILSTTSKSSTSTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVGELTATSPESSTPSST
SVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLV
EGSSTVQQPTETSAPSILSTTSKSSTVTETIGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTV
EELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAET
TLSTTTAAIPSTTFSVVEGSTTVQPSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTG
QTIQPTTTATTTVEELTATSPESSTPSSTSLPVRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVTVESSTEFEVSTQ
GPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTS
ETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVS
TTVTVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAAIPSTTFSVVEGSTTVQPSAPSILSTTSKSSTATETTGVVT
TTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPDSSTPSSTSLPVRQTSPAITSTQPTTV
TKTTAVPVTTLVSTTVTVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILS
TTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVP
LRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAAIPSTTFSVVEG
STTVQPSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTSEEGTTL
ETTKSTISPVVTTNVFYFSTSMETSIAPITKTLPFFTLTIPVTTDSSLFSTTSEFCCFVNETNFHSGDIIYNVTDGF
GWCFTAYCDESCNIVKISTSCVTSPSPSIPPETTTYSSSSTAGTNEPETIISTTNTITEETTPSLPRCDAIDPPRQP
NEAWWSECQKCECDSQTLTVICRPVNCPVSPPITCDKPGQMLITVATVDCCEIQDCTCNSSLCPTPPQQCQPGFTLT
```

```
PVITADACCPEYTCSPMEVCVQNNTVYQPGSNIPSDDPCTQCLCGQDTDPETNLLAPECHSLACMMYCPEETHEYRK
LPGQCCGECVSTECVVVHDNITVNIPVNQTWQPDDKCLTYRCERNNGVSMPVEIKTVCPAFNPEDCIPGTETLDADG
CCPTCTLENKCTVQKNSTFLVSKGCKSSTPVEITSCSGLCGTSSIYSAEANALVHSCSCCQETTTSTKEVTLNCPDG
SNISFSYIYIESCGCKASNCPDQSTNMRRRRRRK
```

```
>danrer_alternative_Muc5.3 | chr7 | obtained by genewise alignment, may not
be reliable
```

```
HNELICSTWGNYHYKTFDGDYFQLPSTCNYILTSHCSSSYEDFNVQLRHELLNNEPTVSKITMKLQGTIIELSRDSL
SVNGQTVTLPYSSAEVFIERVTSYISIKAALGLVAMWNEDDSFMIRLDEKYRNQTCGLCGDFNGIPVYNEFVSGGSS
LSVSDYGNLWKLDGPTETCTEDTMFLNENCGDEYNFCRDIFSSPVFSDCRGLVSIDSFVQVCIQDLCHCNSSSGAFC
LCDTIAEYSRQCVHAGGKPEEWRTEYFCPHPCPETMVYSQCGNPCTDTCTNSERGQVCAQHCIDGCFCPPGSTVYDD
IHQSGCIPYTECSCVLGGKIYASGENYTSNCRDCTCDQGQWDCYIKDCPHSCSVEGGSHITTFDGKAYTFHGDCTYV
LSKHCSGTEFTVLGDLVKCGLTETETCLKAITVALFDGITVINVDSSGNVEVNRIVAQLPLFTSSVSVFKPSSFYII
IEAPLIGLRLEIQLVPIMQLYITANSVYQGETCGLCGNYNNMEADDFTTIGGLREGTALDFANTWKTRASCPDVKRS
FENPCSLSTENEQYAQHWCSLLSDSSSVFAPCHTEVNPEVYKANCIYDSCNCERSEDCMCAALSSYAHACAARGILL
EGWRDAACMKYTETCAATTVYSYQMTSCGRTCRSLSQADGSCTHDHVLVDGCGCAEGKYMNENGDCVEAAACSCYDS
GKVVALGETISKDGGSCICQQGKLSCSGVSNGEECTEPMRFFNCSSASPGSEGAECQKSCNTLDMACISTGCVSGCV
CPSGLVSDGNGGCIDKDQCPCIHNGHTYQSGESIKIDCNTCSCQNRRWTCTTNQCSATCSIYGDGHYRTFDDKRYVF
SGNCEYSLVQDFCNSTSGTFRVITENIPCGSTGTTCSKAIKLFLGSNELRLTDGSFQVVRRDAGEEIPYQMRTMGLY
LVIETKNGLMLIWDRKTTIHIKLGPEYNGRVCGLCGNYDGNANNDFTTRTQAVAVQALDFANSWKLSSCPDATLIQD
PCAHNPYREAWAQRQCSIITSSVFSACHSQVDPSPFYDACVRDACACDSGGDYECFCTAVTAYAQACNEAGACVAWR
SPKICPLFCDYYNPPGECEWHYKPCGAPCMKTCRNPDEQCSNQIPALEGCYPQCPQEQPVFDEDNMKCVKQEECGCF
VDMEHYEVGEQVPTTENCQSCNCSSTGVICEFDINACRCQYNNDTYYYGDIIYHTTDGLGGCITGICSVNGTIERSI
DPCETTTVPTTTSSLITTPTPTTVFVFSTPEQQVYVTTKPSTSLPTTMKKTSESETSTNILTTTMETSIAPITKTLP
TVTDKPWTTLSTQSQSVSTRQTTEEQTTENELTTLETTTSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVT
TLVSTTVTVESSTELEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTST
ETTGVVTTTETVTENELSTSETTMSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRKTSPAIT
SSQPTTLTKTTAVPVTTLVSTTVSVESSTEFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTET
SAPSILSTTSKSSTSTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESST
PSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTPSAPSILSTTSKSSTATETTGVVTTTETVTENELSTSEK
TLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPPITSTQPTTVTKTTAVPVTTLVSTT
VSVESSTKFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETTGVV
TTTETVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTT
VTKTTAVPVTTLVSTTVTVESSTELEQPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTSETTLSTTRPP
KVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVSVESSTE
FEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETTGTTLSTTRPPK
VVTTGQTIEPTTTATTTVEELTPTSPESSTPSSTLVPPTTLTKTTAVPVTTLVSTTVTVESSTEFEVSTQSPSSTTK
PIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTT
RPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSVES
STEFEVSTQGPSSTTKPIKAETTLSTTTAAIPSTTFSVVEGITETTGVVTTTETVTENELSTSETTLSTTRPPKVVT
TGQTIQPTTTATTTVEELTATSPESSTPSSTSLPVRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVTVESSTEFEVS
TQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELS
TSETTMSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTVTKTTAVPVTTL
VSTTVSVESSTEFEQPTETSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPQVVTTGQTIQPT
TTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVTVESSTELEVSTQSPSSTT
KPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTSETTMST
TRPPKPSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTSRPPKVVTTGQTIQPTTTATTTVEELMAT
SPESSTPSSTSLPVRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVTVESSTEFEVSTQGPSSTTKPIKETSAPSILS
TTSKSSTSTETTGVVTTTETVTENELSTSETTMSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVP
LKQTSPTITSTQPTTVTKTTAVPVTTLVSTTVSVESSTEFEVSTQSPSSTTKPIKAETTLSTTTAETPSTTVLLEGS
STVQQPTDTSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTTSQTIQPTTTATTTVEEL
TATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTTTLSTSRPPKVVTTGQTIQPTTTATTTVEE
LTATSPESSTPSSTSLPVRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVTVESSTEFEVSTQGPSSTTKPIKAETTL
STTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTETVTENELSTSETTLSTTRPPKVVTT
GQTIQPTTTATTTVGELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSTQSPSSTTKPI
```

```
KAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETIGVVTTTETVTENELSTSETTLSTTRP
PKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVSTQGPS
STTKPIKAETTLSTTTAAIPSTTFSVVEGSTTVQPSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTLST
TRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSLPVRQTSPAITSTQPTTVTKTTAVPVTTLVSTTVTVE
SSTEFEVSTQGPSSTTKPIKAETTLSTTTAETPSTTVLVEGSSTVQHPTETSAPSILSTTSKSSTSTETTGVVTTTE
TVTENELSTSETTLSTTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKT
TAVPVTTLVSTTVTVESSTEFENELSTSETTLSTTRPPKTNLTSHNIHSTNNSNKDNCCTSHYTCVHNCNCTQGPSS
TTKPIKAETTLSTTTAETPSTTVLVEGSSTVQQPTETSAPSILSTTSKSSTVTETTGVVTTTETVTENELSTSETTL
STTRPPKVVTTGQTIQPTTTATTTVEELTATSPESSTPSSTSVPLRQTSPAITSTQPTTLTKTTAVPVTTLVSTTVS
VESSTEFEVSTQGPSSTTKPIKAETTLSTTTAAIPSTTFSVVEGSTTVQPSAPSILSTTSKSSTVKEGTTLETTKST
ISPVVTTNVFYFSTSMETSIAPITKTLPFFTLTIPNVTDGFGWCFTAYCDESCNIVKISTSCVTSPSPSIPPETTTY
SSSSTAGTNEPETIISTTNTITEETTPSLPRCDAIDPPRQPNEAWWSECQKCECDSQTLTVICRPVNCPVSPPITCD
KPGQMLITVATVDCCEIQDCTCNSSLCPTPPQQCQPGFTLTPVITADACCPEYTCSPMEVCVQNNTVYQPGSNIPSD
DPCTQCLCGQDTDPETNLLAPECHSLACMMYCPEETHEYRKLPGQCCGECVSTECVVVHDNITVNIPVNQTWQPDDK
CLTYRCERNNGVSMPVEIKTVCPAFNPEDCIPGTETLDADGCCPTCTLENKCTVQKNSTFLVSKGCKSSTPVEITSC
SGLCGTSSINRYSAEANALVHSCSCCQETTTSTKEVTLNCPDGSNISFSYIYIESCGCKASNCP
```

Supplementary Addendum S2. Protein sequences of gel-forming secreted mucins used for the construction of the phylogenetic trees. Multiple alignment is in FASTA format and covers only the N-terminal part of the secreted mucins including three VWD domains, before the PTS region.

```
>danrer_Muc2.2_alt_genewise
--STWESFHFKTFDGDIYQYPGLCEYNLVSDCTS--KQFSVDVKRTQ--T
SSGPKISRIVVSINGKMFSISENNVMIDDQ-ITKLPVYESGILIE---EN
TIYTNIRSKTGLTVMWSKDGAVTVELDNKYQNKTCGLCGNFNRIRDEF-I
TGNTEKIGYIEFGNKHKVFNPMEFCENPSETDEKNNWKAICGS-------
------------YQADCEELFRAKAWSSCAMSYHP--YVDACVYDRCFSK
-----SSSLCATLSEYSRQCSRNGGTPSTWRTKDFCAVTCPYNTVHFESG
SPCIDTCSSENTNTQCEEHNIDGCFCPPGTVFDDILNMGCIPREQCHCKH
DRIYKPQEELHEVGQDCVCQNGKWNCKIKPVLGRCSVEEGSHFTTFDGQE
FTFHGDCTYVLSKDSVGSK-FIILGQISPCASLIYDTCLKSIDVIFNSDT
VNRLVISEDGTVKHNAAVLLPYTPASRQFVVFSPSSFYILIDT-NFGLRV
LVQLVPLMQVYITMDKRFQGK------------------------THGLC
GNFNKVLADDMKTPQGLVEGTAVSFANAWEVE-TNC-S-RSDKSVIPCSH
NSVSENYAEHWCYKIMDKSDVFAQCHSKVNPESYYKRCKYASCTCERSED
CLCAVFSSYVHACMIKNVTLEGWRENVCGKYIQNCPAPQRYFYYLKSCQP
TCLSLSSEHTSCKEDFVPVDGCACPDGLYKDQDGQCVPIEKCPCYGMG-K
IAQPGTSIDINGQYCVCVNGVLHCESEKPHIL---ECPSPKVLFNCSTAR
PDQHGLKCKPTCMQKIENC------FSQTCESGCLCPSGLLDDDRGNCVE
RKDCPCPHGGQLYQPETEITVDCNKCTCQSGIWTCTKNK-CPGTCTIYGS
GHYKTFDNQRFGFQGSCSYIAAQDMCG---NRTGIFQVITENIPCGRTGT
TCSKSVFIQFGRVSLHLSDELETQYLME--SGALIR----KIGMYLVIDT
NIGLTVLWDRKTTIHIILQPEYMGEVCGLCGNFNGDGKDDFTTKGGVQTS
NVIEFVDSWKQKTSCPDAAPDFDPCVKNLYREEWAVLKCSIIT--SDTFK
DCQKKVDPTPYYDNCLKDTCACDTGGDCECLCTAVAAYAQACNEADVCVN
WRTPDFCPVYCDYYNKPDECTWHYSPCHPPCYKTCLNPS-GICN----NT
LPNLEGCFL

>danrer_Muc2.1
```

```
--SMWGNFHFKTFDGDVYQFPGMCEYNLVSDCQSLIRQFSIYVKRTE--R
STGPKISRVSITINDIAIELTENQVNVNEA-KVTLPVHVSGILVE---EN
TIYTRLYSKMGITVMWNKDDAVMVELDSKYSNRTCGLCGDFNGIPVYNEF
IQSGRTVGYTEFGNMHRVPNPTHQCEDPFENVDEQNVVDQCEK-------
------------YRADCADLLEDEKWSSCSWVLDPEAYIKACTNDLCNRQ
PEDEDTTALCATLTEYSRQCSHAGGNPPAWRTAKFCNV-QCPYNMVHSES
GSPCMDTCSHKDTNALCEEHNIDGCFCPPG-TVFDDISDT-GCIPAEKCQ
CKHDR-VYSSGEVLRKSEEECYCQEGSWVCMSIPGPGLCAVEEGSHFTTF
DGKEFTFHGDCNYLLSKDCEESK-FIILGQIVPCFTHETDTCLKSIVVLF
DNDKKNPLFIKADGTVQHNAEVSLPYMTAD--FTVFRPSSFHIILQT-VF
-GLQVQVQLVPLMQVYITVDQSFQGK------------------------
TCGLCGNFNKVLADDLKTPQGVVEGTAVSFANAWKAQ-SNCPDRSERMDD
PCSYSSDSEHFAEHWCSKMKDKESLFAQCHATVNPDSYYKRCKYSSCTCE
KSEDCLCTVFSSYARACAAKGVILEGWREVVCDKYIKDCPASQSYSYELQ
SCQRTCVSLASERQSCNVDFVPVDGCACPDGLYQDENNLCVPMEKCPCYH
NG-QKIHPGKSLTISSLSSVCMNGKFHCRSWKTQILG---CPSPKVFFNC
STAGPEEYGLECAKTCSHQEVDC------FSLDCQSGCQCPFGLLDDGRG
HCVQPYNCPCKHDGQYYAPGSVIVKDCNKCICRWGNWLCSQKK-CPGICT
IYGSGHYKTFDQQRFGFRGDCSYIAAQNKCG---NKTGNFHVVTENMPCG
TTGTTCSKAVGILLGRTWLLLSDGTVTATDVG--TGPNIKYNERNVGMYL
VIDAAIGLTVLWDRKTTVRIILQPQHMDDVCGLCGNFNGNGKDDFTTQGN
LPTTNIMEFVDSWKVLSTCPDAQPDFDPCSQTPNRETWAKIQCSIIK--E
VTFKDCHSKVDPNPYYENCVKDSCSCDTGGDCECFCTAVAAYAQACNEAG
VCVVWRTPEICPVYCDYYN-EPGECTWHYSPCHTPCYKTCLNPD-GICN-
---NTLPNLEGCYP

>homsap_MUC2
--STWGNFHYKTFDGDVFRFPGLCDYNFASDCRGSYKEFAVHLKRGPGQA
EAPAGVESILLTIKDDTIYLTRHLAVLNGA-VVSTPHYSPGLLIE---KS
DAYTKVYSRAGLTLMWNREDALMLELDTKFRNHTCGLCGDYNGLQSYSEF
LSDGVLFSPLEFGNMQKINQPDVVCEDPEEEVAPASCSEHR---------
--------------AECERLLTAEAFADCQDLVPLEPYLRACQQDRCRCP
GGD-T--CVCSTVAEFSRQCSHAGGRPGNWRTATLCPK-TCPGNLVYLES
GSPCMDTCSHLEVSSLCEEHRMDGCFCPEG-TVYDDIGDS-GCVPVSQCH
CRLHGHLYTPGQEITNDCEQCVCNAGRWVCKDLPCPGTCALEGGSHITTF
DGKTYTFHGDCYYVLAKGDHNDS-YALLGELAPCGSTDKQTCLKTVVLLA
DKKKNAVVFKS-DGSVLLNELQV-NLPHVTASFSVFRPSSYHIMVSM-AI
GVRLQVQLAPVMQLFVTLDQASQGQ------------------------V
QGLCGNFNGLEGDDFKTASGLVEATGAGFANTWKAQ-STCHDKLDWLDDP
CSLNIESANYAEHWCSLLKKTETPFGRCHSAVDPAEYYKRCKYDTCNCQN
NEDCLCAALSSYARACTAKGVMLWGWREHVCNKDVGSCPNSQVFLYNLTT
CQQTCRSLSEADSHCLEGFAPVDGCGCPDHTFLDEKGRCVPLAKCSCYHR
G-LYLEAGDVVVRQEERCVCRDGRLHCRQIRLIGQS---CTAPKIHMDCS
NLTALATSKPRALSCQTLAAG------YYHTECVSGCVCPDGLMDDGRGG
CVVEKECPCVHNNDLYSSGAKIKVDCNTCTCKRGRWVCTQAV-CHGTCSI
YGSGHYITFDGKYYDFDGHCSYVAVQDYCG-QNSSLGSFSIITENVPCGT
TGVTCSKAIKIFMGRTELKLEDKHRVVIQRD--EGHHVAYTTREVGQYLV
VESSTGIIVIWDKRTTVFIKLAPSYKGTVCGLCGNFDHRSNNDFTTRDHM
VVSSELDFGNSWKEAPTCPDVSTNPEPCSLNPHRRSWAEKQCSILK--SS
VFSICHSKVDPKPFYEACVHDSCSCDTGGDCECFCSAVASYAQECTKEGA
CVFWRTPDLCPIFCDYYN-PPHECEWHYEPCGNRSFETCRTINGIHSN--
--ISVSYLEGCYP

>galgal_Muc2
--STWGNNHFKTFDGDIYQFPGMCEYNFVSDCQDSYREFSVHIQRALNSN
NHPE-IQYILITVTDFTVYLKPKLAVVDGR-IVKTPYYSSGLLIE---SN
DIYTKVYAKLGLILIWNQEDALMVELDSKFGNQTCGLCGDYNGVPIYNEF
```

```
ISGVASYNSITYGNLQKISKPNDECEDPDETRALPSCNEHVSKPNCELRF
PGCMETSLHCEMLREECEKLLTSSAFADCHSRLNLEMYIQACMQDKCACK
GNE-DSFCLCSTISEYSRQCSHAGGRPG----------------------
-------------------------------TVYDDITEN-GCIPVSQCS
CRLHGKGYSPGETITNECEECTCDSGRWTCKDLPCPGTCSVEGGSHIKTF
DGKKYTFHGDCYYVLAKSAVNDT-HALLAELAPCGSSDGQTCLKTVVLLV
DDKKNVVVFRS-DGSVSLNEMTV-NVPHVSASFSVFKPSSYYLIVHT-SF
GLQLQIQLSPVMQLFVTADQSVQDIQGKCPLKQEAKPCISTAVKSYILYM
GGLCGNFNGMEGDDFKTTNGLVEATGSAFANTWKAQ-PTCADQAEKLEDP
CTLSIESANYAEHWCSLLKNSEGPFARCHSVIDPAEYYKRCKYDTCLCKD
NEECLCAALSSYSRACAFKGIILGGWRQSVCSDEVSACPGNQVFLYNLTM
CQQTCRSIADGEKYCLQDFAPVDGCGCPDNTYLDNQDTCVPISKCPCYYK
G-SYLEPGEYVTKDGERCVCRNAKIRCTSVTMRMKSITECPSNKTYFDCN
ASPDWTSRTPVQLRCGTSQTD------LYQGECVSGCVCPEGLFDDGRGG
CVEEEDCPCIHNSDWYNHGQSITVDCNTCTCQKGIWSCTKNV-CYGTCMI
YGSGHYITFDGKFYDFDGSCEYVATQ------------------------
--------------KTELKLENKDYKEIQRD--VGDDVHYQNKTVGLYLV
IEASNGVMLIWDKKTTVFIKLTPDYKGKVCGLCGNFDDKANNDFTSRNGL
QETNALTFGNSWKQSSVCPDVTEEIKPCDLKPHRKSWAEKECSLIL--SE
IFKICHSKVNPSPFYDACVHDACSCDSGGDCDCFCSAVAAYAQECTKAQA
CVFWRSPDICPIFCDYYN-PRNECEWHYEPCGS-NVMTCRMLYNVSTN--
--FSVPYLEGCYP

>homsap_MUC5AC
--STWGSFHYKTFDGDVFRFPGLCNYVFSEHCGAAYEDFNIQLRRSQ--E
SAAPTLSRVLMKVDGVVIQLTKGSVLVNGH-PVLLPFSQSGVLIQ---QS
SSYTKVEARLGLVLMWNHDDSLLLELDTKYANKTCGLCGDFNGMPVVSEL
LSHNTKLTPMEFGNLQKMDDPTEQCQDPVPEPPRNCSTGF----------
--------------GICEELLHGQLFSGCVALVDVGSYLEACRQDLCFCE
DTD-LLSCVCHTLAEYSRQCTHAGGLPQDWRGPDFCPQ-KCPNNMQYHEC
RSPCADTCSNQEHSRACEDHCVAGCFCPEG-TVLDDIGQT-GCVPVSKCA
CVYNGAAYAPGATYSTDCTNCTCSGGRWSCQEVPCPDTCSVLGGAHFSTF
DGKQYTVHGDCSYVLTKPCDSSA-FTVLAELRRCGLTDSETCLKSVTLSL
DGAQTVVVIKA-SGEVFLNQIYT-QLPISAANVTIFRPSTFFIIAQT-SL
GLQLNLQPVPTMQLFMQLAPKLRGQ------------------------T
CGLCGNFNSIQADDFRTLSGVVEATAAAFFNTFKTQ-AACPNIRNSFEDP
CSLSVENEKYAQHWCSQLTDADGPFGRCHAAVKPGTYYSNCVFDTCNCER
SEDCLCAALSSYVHACAAKGVQLGGWRDGVCTKPMTTCPKSMTYHYHVST
CQPTCRSLSEGDITCSVGFIPVDGCICPKGTFLDDTGKCVQASNCPCYHR
G-SMIPNGESVHDSGAICTCTHGKLSCIGG-QAPAP--VCAAPMVFFDCR
NATPGDTGAGCQKSCHTLDMT------CYSPQCVPGCVCPDGLVADGEGG
CITAEDCPCVHNEASYRAGQTIRVGCNTCTCDSRMWRCTDDP-CLATCAV
YGDGHYLTFDGQSYSFNGDCEYTLVQNHCGGKDSTQDSFRVVTENVPCGT
TGTTCSKAIKIFLGGFELKLSHGKVEVIGTD--ESQEVPYTIQQMGIYLV
VDTDIGLVLLWDKKTSIFINLSPEFKGRVCGLCGNFDDIAVNDFATRSRS
VVGDVLEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQKQCSILH--GP
TFAACHAHVEPARYYEACVNDACACDSGGDCECFCTAVAAYAQACHEVGL
CVCLRTPSICPLFCDYYN-PEGQCEWHYQPCGVPCLRTCRNPRGDCL---
--RDVRGLEGCYP

>homsap_MUC5B
--STWGDFHYKTFDGDVFRFPGLCNYVFSEHCRAAYEDFNVQLRRGL--V
GSRPVVTRVVIKAQGLVLEASNGSVLINGQ-REELPYSRTGLLVE---QS
GDYIKVSIRLVLTFLWNGEDSALLELDPKYANQTCGLCGDFNGLPAFNEF
YAHNARLTPLQFGNLQKLDGPTEQCPDPLPLPAGNCTDEE----------
--------------GICHRTLLGPAFAECHALVDSTAYLAACAQDLCRCP
------TCPCATFVEYSRQCAHAGGQPRNWRCPELCPR-TCPLNMQHQEC
```

```
GSPCTDTCSNPQRAQLCEDHCVDGCFCPPGSTVLDDITHS-GCLPLGQCP
CTHGGRTYSPGTSFNTTCSSCTCSGGLWQCQDLPCPGTCSVQGGAHISTY
DEKLYDLHGDCSYVLSKKCADSS-FTVLAELRKCGLTDNENCLKAVTLSL
DGGDTAIRVQA-DGGVFLNSIYT-QLPLSAANITLFTPSSFFIVVQT-GL
GLQLLVQLVPLMQVFVRLDPAHQGQ------------------------M
CGLCGNFNQNQADDFTALSGVVEATGAAFANTWKAQ-AACANARNSFEDP
CSLSVENENYARHWCSRLTDPNSAFSRCHSIINPKPFHSNCMFDTCNCER
SEDCLCAALSSYVHACAAKGVQLSDWRDGVCTKYMQNCPKSQRYAYVVDA
CQPTCRGLSEADVTCSVSFVPVDGCTCPAGTFLNDAGACVPAQECPCYAH
G-TVLAPGEVVHDEGAVCSCTGGKLSCLGASLQKST--GCAAPMVYLDCS
NSSAGTPGAECLRSCHTLDVG------CFSTHCVSGCVCPPGLVSDGSGG
CIAEEDCPCVHNEATYKPGETIRVDCNTCTCRNRRWECSHRL-CLGTCVA
YGDGHFITFDGDRYSFEGSCEYILAQDYCG-DNTTHGTFRIVTENIPCGT
TGTTCSKAIKLFVESYELILQEGTFKAVARG--PGGDPPYKIRYMGIFLV
IET-HGMAVSWDRKTSVFIRLHQDYKGRVCGLCGNFDDNAINDFATRSRS
VVGDALEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQKQCSILH--GP
TFAACRSQVDSTKYYEACVNDACACDSGGDCECFCTAVAAYAQACHDAGL
CVSWRTPDTCPLFCDFYN-PHGGCEWHYQPCGAPCLKTCRNPSGHCL---
--VDLPGLEGCYP

>galgal_Muc5AC
--STWGNFHFKNFDGDIFYFPGVCNYIFASNCKSPYEDFNIQIRRTM--V
ENATVITHVIMRLEGAVVELTRTSVLVNGK-QIQMPYSHMGVLIE---RS
NSYIKVSAKLGLTFLWNEEDALLVELDKKYANQTCGLCGDFNGIPTNNEF
ISQNAKLTAVQFGNRQKIDGPTEQCDDPIPSAVLMNCSSEF---------
-------------ASICKTVLTSEAFTSCNALVDVQDYIDTCIQDLCHCD
TSM-ADFCMCNTFAEYSRQCAHAGGQPLNWRTSELC--------------
------------------------------RMVFDDINGA-GCIPRKECH
CTYEGETYAPGTSFSSKCRSCTCAGGEWTCVTQSCPATCSIEGGSHITTY
DEKYYSFYGDCSYVLTKLCDSND-FTVLGDIHKCGLTDTETCLKGVAISL
SGGQTNIVIQP-SGSVFVNMIYT-QLPFSAANVTIFRPSSFFIILQT-TF
GLQLQIQIVPLMQLFINLDPSQKGQ------------------------T
CGLCGNFNDIQTDDFKTTSGVIEGTSAAFGNTWKTR-ADCPDAKITFENP
CTLSIENDKYAQHWCGLLSDTTGPFAECHSTVNPEVYQKNCMFDTCNCEN
TEDCMCAALSSYVRACAAKGILLNGWRNKVCTKYTNTCPKSLEYTYNVDS
CQPTCRSLSEPDVTCNIKFVPVDGCTCVNGTYMDESGKCVPAASCPCYYK
G-MPLSTGEVIHDNGVVCTCSYGKLSCIGE--KPPP--VCAPPMVYVDCG
NATAGVTGAECQKSCQTLDME------CYKTHCVSGCICPPNYVSDGKGG
CIAAEECPCVHNGDSYRPGESIRVGCNNCTCRNRKWHCSEEP-CLETCSV
YGDGHYTTFDGKRFDFEGDCEYVLIQNYCGRKGLNQGTFRVITENIPCGT
TGTTCSKSIKVFLGNYELVLSDGHSDVIQRS--PGGEMPFQIRSMGIYMV
VDTTVGLILMWDKKTSIFIKLSPSFQGHVCGLCGNYDGNGNNDFTTRSQS
VVGNVLEFANSWKVSSTCPNANQTKPCTANPYRKSWAQKQCSIIT--SEV
FAKCHSQVEPNEYYQACVDDACACDTGGDCECFCTAVAAYAQACNELDIC
ISWRTPSICPLFCDYYN-PQGECEWHYKPCGAPCMKTCNNPSGKCL----
-HELRGLEGCYP

>galgal_Muc5B
--CTWGNFHFKTFDGDIFRYPGLCNYVFASHCNAPYEEFNIQIRRTL--V
ENILTIHRIAIKLEGVAIELTKDVVLINSN-RVQLPYSQAGITIE---RS
SIYVKVVSKTDIVFMWNQRDSILLELHEKYANQTCGLCGDFNGIPVYNEF
YSNNIRMTALQFGNMQKMDGPTEYCEDPTPSSLDN---------------
--------------------------------------------------
--------------------------------------------------
--------------------------------------------------
CTYN----------------------------------------------
---FH-------------CEDEI-FTILVELRKCGMTNTETCLKTVTLSM
```

```
NKGQTLIEVKT-DGGVFMNSIYT-QLPISAANVTIFRPSSFFVIMHT-NI
GIQVEIQITPIMQVFVRLDPVFKDQ------------------------T
CGLCGNFNNVQTDDFKAISGVVEGTAAAFANTWKTQ-ASCPDIQSSFENP
CALSIDNEKYAKHWCGLLTNSEGPFADCHYLLNPSTYYTNCMFDTCNCQN
SEGCLCAALSSYVRACAARGIQINGWRTNVCSKYTTSCPKSLSYSYTINS
CQPTCRSLSEPDVTCKIKFVPVDGCTCVNGTYMDESGKCVPANECPCYYR
G-SPIPLGEVVHENGLVCT-------------------------------
-------------------------------QCVSGCVCPSGLVLDGKGG
CIPEEECPCIHNGAMYQPGEKINADCNTCVCKDRMWECTKKQ-CLGTCAV
YGDGHYTTFDDQRFSFSGNCEYTLVQNHCGKSDTVNGTFQVITENIPCGS
TGTTCSKSIKVFLEHYELVLSEERVSVIKRG--QDGEVPFTVRYIGMYTV
IDTTSGLILMWDKKTSIFIKLSPEFKGQICGLCGNYDGNGINDFTTRSLS
VVENVLEFGNSWKVSSTCPDASFVKDPCSTNPYRKSWSEKQCSIIN--SN
VFAACHSQVEPAKYYQACVTDACACDSGGDCDCFCTAVAAYAQACNEVGV
CIAWRTPSIC----------------------------------------
---------RCYP
>danrer_Muc5.3
--STWGNYHYKTFDGDYFQLPSTCNYILTSHCSSSYEDFNVQLRHEL--L
NNEPTVSKITMKLQGTIIELSRDSLSVNGQ-TVTLPYSSAEVFIE---RV
TSYISIKAALGLVAMWNEDDSFMIRLDEKYRNQTCGLCGDFNGIPVYNEF
VSGGSSLSVSDYGNLWKLDGPTETCTEDTMFLNENCGDEN----------
---------------FCRDIFSSPVFSDCRGLVSIDSFVQVCIQDLCHCN
SSS-GAFCLCDTIAEYSRQCVHAGGKPEEWRTEYFCPH-PCPETMVYSQC
GNPCTDTCTNSERGQVCAQHCIDGCFCPPG-TVYDDIHQS-GCIPYTECS
CVLGGKIYASGENYTSNCRDCTCDQGQWDCYIKDCPHSCSVEGGSHITTF
DGKAYTFHGDCTYVLSKHCSGTE-FTVLGDLVKCGLTETETCLKAITVAL
FDGITVINVDS-SGNVEVNRIVA-QLPLFTSSVSVFKPSSFYIIIEAPLI
GLRLEIQLVPIMQLYITANSVYQGE------------------------T
CGLCGNYNNMEADDFTTIGGLREGTALDFANTWKTR-ASCPDVKRSFENP
CSLSTENEQYAQHWCSLLSDSSSVFAPCHTEVNPEVYKANCIYDSCNCER
SEDCMCAALSSYAHACAARGILLEGWRDAACMKYTETCAATTVYSYQMTS
CGRTCRSLSQADGSCTHDHVLVDGCGCAEGKYMNENGDCVEAAACSCYDS
G-KVVALGETISKDGGSCICQQGKLSCSG--VSNGE--ECTEPMRFFNCS
SASPGSEGAECQKSCNTLDMA------CISTGCVSGCVCPSGLVSDGNGG
CIDKDQCPCIHNGHTYQSGESIKIDCNTCSCQNRRWTCTTNQ-CSATCSI
YGDGHYRTFDDKRYVFSGNCEYSLVQDFCN--STSGTFRVITENIPCGST
GTTCSKAIKLFLGSNELRLTDGSFQVVRRD--AGEEIPYQMRTMGLYLVI
ETKNGLMLIWDRKTTIHIKLGPEYNGRVCGLCGNYDGNANNDFTTRTQAV
AVQALDFANSWKLS-SCPDATLIQDPCAHNPYREAWAQRQCSIIT--SSV
FSACHSQVDPSPFYDACVRDACACDSGGDYECFCTAVTAYAQACNEAGAC
VAWRSPKICPLFCDYYN-PPGECEWHYKPCGAPCMKTCRNPDEQCS----
-NQIPALEGCYP

>danrer_Muc5.1
--STWGNFHFKTFDGHFFQLPDTCNYVLAVMCDAASSDFNIQMQRET--V
NGSISFSTVLIKLDGTVIKVTDSDITMGEE-TVTVPTYKNGIKIE---GS
PTSFKISNKHGVTVFWEEDNSLSIELPEKYQGQTCGLCGDFNGN-LADDI
TDNGP--------ATWKVSISTEICEEVTLPSTGPCDELSEQ--------
-------------ASFCEEYLISPGFSGCYDVMDMRIFQKACVSDLCQCY
GN---HDCLCNTLTEISRQCTHAGGQPGTWRTEQLCPK-TCPLNLQYLEC
GGPCRNTCSAPTANLMCTDHCVDGCFCPEG-TVEDDIGQS-GCVPVNECP
CVHDGTVYKSGESYQQACKSCTCDAGHWTCTYLDCPGTCSVVGGSHVTTF
DGKSFTFSGNCDYILTKHSADSD-FAVVGNLAKCEASRTDTCLHSVTLVI
SG--TTISFSS-SGDVTLNNINSKELPFSIGPVSIFQPSSSFIIVDL--T
SVRLEIQLAPVMQLYIVASTEEKGK------------------------M
TGLCGNYNDVQTDDFKTESGIIEGTPTSFVNFWKQN---CPDLEITFDNP
CSLNMDTEKLAKDWCSRLTNPNGTFSECHSEICPEMYYQRCVYDTCKCSD
```

```
IRKCICAAVSTYAHACAARGIILHDWMD--TEPCESECPENMKHSYEVTS
CGSTCRSLSGQENTCQGSFTPVHGCVCSEGTYLNEEGKCVHAGMCPCYYG
N-QVIQPSAEFNKEGAKCTCNNGKLDCSSH-----E--DCVAPMTFFKCS
--HPGEKGTECQRTCDKQDPNN-----CVSTGCVSGCMCPEGLLADGKGG
CVMRENCPCTHNGVTSSPGDQVQQDCNTCTCTNGMWICTQKA-CYGTCTI
YGEGHFRTFDGNRYSFHGDCEHTLAHDYCD-GNPSP-SFRLVTENIPCAT
ANSICSKSINLFFGRYEMILSEDDGVKVVEG--NGTEYKYQINYGGIYVV
IEIEGLLNLIWDEKTSVMIQLHPKLKGNVCGLCGNFDGNANNDFMKHNGE
VVTDPEEFGNSWKVDSNCPDVTNVMHPCEAHPHRNSWAIKQCRIIT--SP
VFTDCHSLVDSGPYYDACVRDTCACDSGGDCDCLCTAVGAYAAECRKRGA
CVAWRTPHFCPLFCDYYN-APGECQWHYKTCGSTCMKTCRNPLGTCS---
--DQIPPLEGCFL

>danrer_Muc5.2
--STWGNFHFKTFDGHFFQLPDTCNYVLAVMCDAAISDFNIQMQRDT--V
NGSISFGSVIVKLDGTVIKITDSGIMMDDQ-VVSVPINQKGIKIE---GS
PTSIKIS-RYGMTIFWEEDNSIAIELAEKYKGLTCGLCGNYNGD-KDDDM
PESGP--------ATWKISTPTESCEDVILPPKDQCDQN-----------
-------------TMVCQQYLSSPGFSGCYDVMDMKIFEKACVSDMCQCY
GS---HDCLCNTLTEISRQCTHAGGQPGTWRTEQLCPK-MCPINLQYMEC
GGPCKSTCSDPTAHLMCKDHCVDGCFCPEG-TVEDDIGQG-GCVPVNECP
CVHDGTVYKSGESYQQACKKCFCAAGHWTCTYLDCPGTCSVVGGSHVTTF
DGKSFTFSGNCDYILTKHSNDSD-FAVVGNLAKCEPARTDTCLNSVTLVI
SG--TTIGFTS-DGVVTLNGNSPFNLPAVIGPVSIFQPSLSFIIADL--N
SLRLEIQLAPVMQLYIVASTEEKGK------------------------M
TGLCGNYNDVQSDDFKTDLGITEGTAISFANFWKKTPYSCPDLENTFDNP
CSLSVDTEKLAKDWCSRLTNQSGAFSACHSEICPKIYYERCVYDTCKCAD
IRKCICAAVSTYAHACAARGIILQGWMD--SDPCESECSENMKHSYGMTG
CGLTCRSLSGQENTCQGSFTPVDGCVCSEGTYLNEEGICVPADQCPCYSG
D-QVTKPSEVSHVDGLTCTCKLGKLHCSNP-----E--TCVAPMVLFKCS
NYEPGKKGTECQRTCQKQDINN-----CVSTGCVSGCMCPDNLLADGQGG
CVEREKCPCVHNGVTYSSGEQVQEDCNTCTCRNGMWDCTEKE-CYGTCTI
YGEGHFMTFDGKKYSFHGDCEQILVHDYCN-TDQSPYSLRLVTENIPCGT
SGTICSKSINLFFGRYKLILSEEKEIQVVES--NGTDYQYQIHTAGIYNV
IEVKGLLNLIWDSKTSVMLQLHPKFKGKVCGLCGNFDGNANNDFMKHDGE
EVTDPVVFGNSWKSNPGCPDVSNVMNPCEKNPHRSAWAIKQCSIIT--SP
VFSDCHSRVDSGPYYDACVKDTCACDSGGDCDCFCTAVAAYAAECRKTGA
CVAWRSPSICPLFCDYYNNPPSECEWHYKTCGSTCMKTCRNPSGTCS---
--NQIPLLEGCFP

>homsap_MUC6
--STWGAGHFSTFDHHVYDFSGTCNYIFAATCKDAFPTFSVQLRRGP---
--DGSISRIIVELGASVVTVSEAIISVKDIGVISLPYTSNGLQITPFGQS
VRLVAKQLELELEVVWGPDSHLMVLVERKYMGQMCGLCGNFDGKVTNEFV
SEEGKFLEPHKFAALQKLDDPGEICTFQDIPSTHVRQAQH----------
-------------ARICTQLLT---LVAPECSVSKEPFVLSCQADVAAAP
QPG-PQNSSCATLSEYSRQCSMVGQPVRRWRSPGLCSVGQCPANQVYQEC
GSACVKTCSNPQHS--CSSSCTFGCFCPEG-TVLNDLSNNHTCVPVTQCP
CVLHGAMYAPGEVTIAACQTCRCTLGRWVCTERPCPGHCSLEGGSFVTTF
DARPYRFHGTCTYILLQSPQLPEDGALMAVYDKSGVSHSETSLVAVVYLS
RQDKIVISQDE----VVTNNGEAKWLPYKTRNITVFRQTSTHLQMAT-SF
GLELVVQLRPIFQAYVTVGPQFRGQ------------------------T
RGLCGNFNGDTTDDFTTSMGIAEGTASLFVDSWRAG--NCPAALERETDP
CSMSQLNKVCAETHCSMLLRTGTVFERCHATVNPAPFYKRCVYQACNYEE
TFPHICAALGDYVHACSLRGVLLWGWRSS-VDNCTIPCTGNTTFSYNSQA
CERTCLSLSDRATECHHSAVPVDGCNCPDGTYLNQKGECVRKAQCPCILE
GYKFILAEQSTVINGITCHCINGRLSCPQR--PQMFLASCQAPKTFKSCS
```

```
QSSENKFGAACAPTCQMLATG----VACVPTKCEPGCVCAEGLYENADGQ
CVPPEECPCEFSGVSYPGGAELHTDCRTCSCSRGRWACQQGTHCPSTCTL
YGEGHVITFDGQRFVFDGNCEYILATDVCG-VNDSQPTFKILTENVICGN
SGVTCSRAIKIFLGGLSVVLADRNYTVTGEEPHVQLGVTPGALSLVVDIS
IPGRYNLTLIWNRHMTILIRIARASQDPLCGLCGNFNGNMKDDFETRSRY
VASSELELVNSWKESPLCGDVSFVTDPCSLNAFRRSWAERKCSVIN--SQ
TFATCHSKVYHLPYYEACVRDACGCDSGGDCECLCDAVAAYAQACLDKGV
CVDWRTPAFCPIYCGFYNTHTQDGHG-EYQYTQEANCTWHYQPCLCPSQP
QSVPGSNIEGCY-

>galgal_Muc6
--STWGGGHFSTFDKYQYDFTGTCNYIFATVCDESSPDFNIQFRRGL---
--DKKIARIIIELGPSVIIVEKDSISVRSVGVIKLPYASNGIQIAPYGRS
VRLVAKLMEMELVVMWNNEDYLMVLTEKKYMGKTCGMCGNYDGYELNDFV
SE-GKLLDTYKFAALQKMDDPSEICLSEEISIPAIPHKKY----------
-------------AVICSQLLN---LVSPTCSVPKDGFVTRCQLDMQDCS
EPG-QKNCTCSTLSEYSRQCAMSHQVVFNWRTENFCSVGKCSANQIYEEC
GSPCIKTCSNPEYS--CSSHCTYGCFCPEG-TVLDDISKNRTCVHLEQCP
CTLNGETYAPGDTMKAACRTCKCTMGQWNCKELPCPGRCSLEGGSFVTTF
DSRSYRFHGVCTYILMKSSSLPHNGTLMAIYEKSGYSHSETSLSAIIYLS
TKDKIVISQNE----LLTDDDELKRLPYKSGDITIFKQSSMFIQMHT-EF
GLELVVQTSPVFQAYVKVSAQFQGR------------------------T
LGLCGNYNGDTTDDFMTSMDITEGTASLFVDSWRAG--NCLPAMERETDP
CALSQLNKISAETHCSILTKKGTVFETCHAVVNPTPFYKRCVYQACNYEE
TFPYICSALGSYARTCSSMGLILENWRNS-MDNCTITCTGNQTFSYNTQA
CERTCLSLSNPTLECHPTDIPIEGCNCPKGMYLNHKNECVRKSHCPCYLE
DRKYILPDQSTMT------------------------ESCKAPKKYISCS
DSLENKYGATCAPTCQMLATG----IECIPTKCESGCVCADGLYENLDGR
CVPPEECPCEYGGLSYGKGEQIQTECEICTCRKGKWKCVQKSRCSSTCNL
YGEGHITTFDGQRFVFDGNCEYILAMN-----------------------
---------------LTIILRDETYSISGKNLQVKYNVKKNALHLMFDII
IPGKYNMTLIWNKHMNFFIKISRETQETICGLCGNYNGNMKDDFETRSKY
VASNELEFVNSWKENPLCGDVYFVVDPCSKNPYRKAWAEKTCSIIN--SQ
VFSACHNKVNRMPYYEACVRDSCGCDIGGDCECMCDAIAVYAMACLDKGI
CIDWRTPEFCPVYCEYYNSHRKTGSGGAYSYGSSVNCTWHYRPCNCPN--
QYYKYVNIEGCY-
```

Supplementary Addendum S3. Sequences of transcripts obtained by RT-PCR and sequencing to verify the expression of mucin genes. The primers used for amplification of each transcript are also shown.

muc5.1

F1- IJ1 – tgacatgggttgaaagcaaa
R1 – IJ4 – tggtggtgttgaaaatgtga

```
TGACATGGGTTGAAAGCAAACTAAGAGACATTATAAATAGACAACTGTTAGCAGATGACTCACATTAGCTCTTGAGG
ATGGGTTTCGACTCTGAGGGGACGGTCAGGATGCCCCAGATGTGGATGCTGAGATGGGTCTTCCTACTGGCCGGGCT
GCAGTCCATACAAGCAGGCTTTATGGGAGATTACAGAGATATGGAAAATCCTATGACTCCTATGTGGCCAACAACAC
CAGCTATCTCTATGACAAGTGTGCCAGTAATTACAGTTGAACCCAACCCCGATCATCGGTCCACGATCTGCAGCACA
TGGGGAAATTTCCACTTCAAGACGTTCGATGGTCATTTCTTTCAACTGCCAGACACATGCAACTATGTTCTGGCGGT
GATGTGTGATGCCGCATCATCTGATTTTAACATCCAGATGCAAAGGGAGACTTTAAATGGCTCAATCTCATTTAGCA
CAGTCTTAATAAAGCTGGACGGGACAGTTATCAAAGTCACCGACAGCGACATCACTATGGGTGAAGAAACGGTCACG
GTTCCTACTTATAAAAATGGAATTAAAATTGAGGGAAGTCCAACCAGTTTCAAGATCTCCAACAAACATGGAGTGAC
TGTCTTTTGGGAGGAAGACAACTCTCTGTCGATTGAACTTCCTGAGAAGTATCAAGGCCAGACCTGTGGACTCTGCG
```

GAGACTTTAATGGCAACTTGGCTGATGATATCACTGATAATGGTCCGGCTACATGGAAGGTGTCCATATCAACTGAA
ATCTGTGAGGAGGTCACTCTTCCATCTACTGGTCCGTGTGACGAGCTATCAGAGCAGGCAAGCTTTTGCGAGGAGTA
CCTCATCAGTCCAGGGTTTAGTGGCTGTTATGATGTGATGGACATGAGAATATTTCAGAAAGCCTGTGTGAGCGACC
TGTGTCAATGTTATGGCAACCATGACTGTCTGTGCAACACACTAACAGAGATCTCACGACAATGCACACATGCTGGA
GGACAACCAGGAACATGGAGAACAGAACAGCTCTGCCCAAAGACGTGTCCCCTAAACCTGCAGTACCTGGAGTGTGG
CGGTCCATGCAGGAACACCTGCTCAGGCCCAACTGCAAATCTAATGTGTACCGATCATTGTGTGGACGGCTGCTTCT
GCCCTGAAGGCACCGTTGAAGATGACATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNGAGAGTCTTACCAACAAGCTTGCAAATCATGTACATGTGATGCTGGACACTGGACCTG
CACATACCTGGACTGTCCCGGCACATGCTCTGTTGTTGGAGGTTCTCACGTCACCACTTTTGATGGAAAGAGCTTCA
CCTTCAGCGGGAACTGCGACTACATCCTCACTAAGCACTCTGCAGACAGTGACTTTGCTGTTGTGGGAAATCTGGCA
AAGTGTGAAGCATCCCGAACAGACACCTGCTTACACTCAGTTACCCTTGTCATCTCTGGGACCACTATTAGTTTTTC
CTCCAGTGGAGATGTGACATTAAATAACATCAATTCAAAAGAGCTTCCTTTCAGCATTGGTCCTGTGAGTATCTTCC
AGCCCTCTTCCTCCTTCATTATTGTTGATTTGACGAGCGTTCGTCTGGAGATTCAGCTGGCTCCAGTCATGCAGTTG
TACATTGTAGCCAGCACTGAAGAGAAAGGAAAAATGACTGGTCTGTGTGGGAACTATAATGATGTCCAAACAGATGA
TTTTAAAACTGAATCAGGCATTATCGAGGGCACCCCAACATCCTTTGTCAACTTTTGGAAACAAAACTGTCCAGATC
TTGAAATCACATTTGACAACCCATGCAGCTTGAACATGGACACAGAGAAACTTGCTAAAGACTGGTGCTCCCGCCTT
ACAAATCCAAATGGCACCTTTTCCGAGTGTCACTCAGAAATATGTCCTGAGATGTACTACCAAAGGTGCGTCTATGA
CACCTGCAAGTGTTCAGATATTAGGAAGTGCATCTGTGCAGCCGTGTCCACCTATGCCCACGCATGTGCCGCCAGAG
GAATCATTCTCCACGACTGGATGGATACAGACCCTTGTGAATCTGAGTGTCCGGAAAACATGAAGCACTCCTACGAA
GTGACCAGCTGTGGCAGCACCTGCCGCTCTCTCAGTGGACAAGAAAACACCTGCCAAGGGTCCTTCACACCTGTACA
TGGCTGTGTTTGTTCTGAGGGAACCTACCTTAACGAAGAAGGCAAATGTGTACATGCTGGAATGTGTCCTTGTTATT
ATGGCAACCAGGTCATACAGCCATCCGCAGAGTTCAACAAGGAAGGTGCTAAATGCACCTGTAACAATGGCAAACTG
GATTGCTCCAGTCATGAAGACTGCGTGGCTCCAATGACATTCTTCAAATGCTCACATCCTGGAGAAAAAGGCACAGA
ATGTCAGAGGACATGCGACAAACAAGACCCAAACAATTGCGTGAGCACAGGGTGTGTATCAGGGTGTATGTGTCCTG
AGGGCCTTTTGGCTGATGGAAATGGTGGATGTGTGATGAGGGAAAACTGTCCATGCACCCATAACGGTGTCGCAAGT
TCACCTGGGGATCAAGTTCAACAGGACTGTAACACCTGCACATGCACAAACGGTATGTGGATTTGTACACAGAAGGC
ATGTTACGGCACCTGTACCATCTACGGTGAAGGTCATTTCAGGACTTTTGATGGCAACAGATACTCCTTCCATGGAG
ATTGTGAACACACCTTAGCCCACGATTATTGTGATGGAAACCCCAGTCCCTCTTTCCGCCTTGTCACCGAAAACATT
CCCTGTGCCACAGCCAACTCCATTTGCTCCAAGTCTATTAATCTTTTCTTCGGGAGATATGAGATGATTTTATCAGA
GGATGATGGCGTGAAAGTTGTTGAAGGCAACGGCACAGAATACAAATATCAAATCAATTATGGAGGCATATATGTGG
TCATAGAGATCGAAGGTCTTCTGAACCTAATCTGGGATGAGAAAACAAGTGTGATGATCCAGCTCCATCCCAAATTA
AAGGGCAATGTGTGTGGACTGTGTGGAAACTTTGATGGCAATGCAAACAATGACTTCATGAAGCACAATGGGGAAGT
GGTGACAGATCCAGAGGAGTTTGGCAATAGCTGGACGGTGGATTCAAATTGTCCTGATGTGACAAATGTGATGCATC
CTTGTGAAGCACATCCACATAGGAATTCCTGGGCTATCAAGCAATGCAGA

F3 - IJ7 – catccctaggcacatccact
R3 - IJ63 – tgttttgcgtcgcttaagaa

CATCCCTAGGCACATCCACTGTTACACCAACTACTGGAACTTCGGCACCAACTTCTACAGAAACACCATCCACAACA
ACATCCATTACTTCACCAACGACTGAAACATCAACACCAACTTCTACAGAAACACCATCCACAATCACATCTACTGT
TACACCAACTACTGAAACATCAACACCTACTTCAACAGAAACTACATCTACAACAATTCCAGTAACAACTATACCAA
CAGAAACTCCATCTACAGTTGTTACAGTAACAACTTTATCAACAGAAACTCCATCTACAACAGTTTCAGTAACAACT
ATACCAACAGAAACTCCATCTACAATTGTTACAGTAACAACTATATCAACAGAAACTCCATCTACAACTGTTACAGT
GACACCTATATCAACAAAATCTCCATCTACAATTGTTACAGCAACACCCAAATCAACAAAAACACCATCTATAACTA
TTGGATCTACAACAACTATGTCAACAGAAAAACCATCTACATCTAGCAAATCTACAACAACATCCACATCAACAAAG
CCTGTTACTACTACAACAAAGTTCACTGTGCATACAGGTCCAGTTCGCACCACTCCAAATCCAGAACAACCATCAAC
TCAATGTAACTGCACATATTTAAATGCAATATTCTTACCAGGGTCATCAATATACAATAAAACAGATAAAGCAGGTT
GGTGCTTTTCCGCTCACTGCAACTCCTCGTGCAGTATTATCATACAGGCTGAGCCATGTGCAACAACAACACCACCA
ACAACAACACCAACAACGACACCAACAACGACGACATCCAAACCGAACAAAAATGATTGCACACAGGTCCAAAGAAA
GAATATGGAGTCCTGGATTGAAAACTGCAAAAACTACACGTGTAAAAATGGCAAGGTAACCTCAGTGCCTCTACAGT
GTGACAAATCTGACAAGGTCATGCCTACATGTACCAATGGAATTAAAGCTAAGAAGGTGTCTTATAACAACGGCTGC
TGCTCTAGATATGAATGCGAATGTAAATGCAGTGGTTGGGGTGACCCACATTACCAAATGTTTGATGGAACATACTA
TGACTTCCAAGGAAATTGCACTTACGTCTTGTTCAAAGAAATCGTCCCAAGATATAACATCAGTGCCTATGTTAAAA
ACTACTTATGTGATCTTCAAAATAATCTTGCCTGCCCAGAGTATGTGATTGTGAACTACAAATCATATAGCATTAAG
CTGACAAGCAACACCAAAGAGATTGAGGTCTATGTTAATGATGATTTGAAACAACTAACTTACGTCAATAACAACTT
CTTCATCACCACCTCTGGCATTGGAGTCATTTTGAACATCACTGAAATCAATGTCGAGATTACTGTGAATCATCAAG

```
GCTTTGTGATCAATCTCCCATTCTCCTACTTCAGTGGTAATACAGAGGGGCAGTGTGGGGTCTGCGACAACAATCGT
ACAAATGACTGCAGACGTCCTGATGGACAAATCGACAAATCGTGTGAGAACATGGCACAGCTTTGGATGGTTCCCCC
AGGATGCCAGCCTCCTAAACCTATACCTAAGCCTACACCTCCTCCAGTTATCTGCAACCCCAGTATTTGTGAATACA
TCAAGGGCGACATGTTTAAGGAGTGCCATGATCATATTCCTTACGAAAGTTATTACGAGTCATGCAAATTTGACGTG
TGCCGCATGGGAAATCGCTCAATTGGTTGCACCAGTCTTGAGGCGTATGCTGCATTGTGTGGAAAAGAAGGCATCTG
TGTGGACTGGAGAACCTCAGATGACCTCAAGCGAACATGTGCACACAATTGCCCCAGTCACAAAGTCTACAATGCAT
GTGGACCTAAAGTGGAAAAAACCTGCAGTACCAGGTACAATGACTTGTTTGTGGAGAAAGACTGTCAAGGCGCAAGC
TGTCATCAAACATATATGGAAGGTTGTTTCTGTCCTGATAACACCTATCTAGTTAGTTCAACGACAGATTACTGTAC
GGCTTACTGCGACTGCATCGGCCCTGATGGATTACCTCGAGAGCCTGGAGACACATGGACCATGGACTGTTACACAT
ACACCTGCTCCAATGAGACCTTTGGGATCACCAAAGAACTTGTTGGGTGTCCGACTATAGTACCCTGTGGACCTGAG
CACAAACTAATCATTAACAACTGCTGTCCAACCTGCGTGTGTGATCTGGAGGTCTGTCTCCACAAGAAATGTGATGT
GGGATTCGAGTTAGCGGAACACGCAAATGAAGGCAGTTGCTGTCCACTCTGTGTGCCCAAAGACGTCTGTGTGTACA
ATAACACCGAGTACCAGCCTGGATCCACAATCTACCCAGACCCTTGTGTCGAATGTAACTGTGACATGGAGAAAGAC
CCAGAGACACAACTACACACCGTCTCCTGCGTTTCTAAAATCTGTGAACCTTGCCTTGAGGGCTTTGAGCTCGTGGA
ACAAAACGGAGAGTGTTGTGGAACATGCAAGCCAAATGCCTGTATTTATGTGTCAGATAACATCACACATATTCTCA
GTGAGGGAGAGAGTTACAATTATAAATGTGAAAGTGTGACCTGTCAGCAAGTCAATGGGACGTACATAATAGAGAGG
ACCATACCAACCTGTCCCGAAATCAATCCAGATCAATGTGTGCCTGGAACACTGGGACTCGATGAAGATGGATGCTG
CAACACCTGTGAACTCAAGAACTGTGTTCGTGTGAAGAACATGACAGATGTCACAGTGAATGACTGCAAGTCCATCA
ACCCCATTGAAGTCACATCCTGCAGTGGAAACTGCGACACAGAATCAATGTATTCAATGGGCGCAAACACCATGATG
CACAGCTGTTCCTGCTGTAAGGAGACGAAAACCAGCGTTAAAAAGGTGACGCTGAAGTGTGCGGACGGCAGTGAAAT
CCCTCATGACTACGTCTACATAGAGAGCTGCAGATGCACTCCTATTACATGCGAAAACCAGAAAACCAATGGATAAA
TGATCATGTTCACATTGGTTTAATGGCTACAATTCCAGATATTACTGGAAAAATATGTATCTTTAGTGAATGACTGT
ACCATTCAGCTCTTTGTTCTATAACTACTGTCACTGAGGGAATTTTCCGGATATTTTGAATTGTTTTGTTTGTTCTT
AAGCGACGCAAAACA
```

muc5.2

F1 - IJ68 – ctgagatgggtcatcctgct
R1 - IJ71 – gtctcgggtgatgaaggtgt

```
ATGGACTGTGGGACGGTGATGTTGTCCCACATGCGGATGCTGAGATGGGTCATCCTGCTGGCAGGGATGCAGTCTAT
ACGAGCAGACTTCATGGGAGATTATGGAGACGGTATGAAGGACCCTATGGACTTTACTTGGCTGCCAACAACAACTG
TGTCCCCCACTATGACGACTGTGATGGTAACAAAAGTTGAACCCAAACCTGATCATCAGTCCCGAATCTGCAGCACA
TGGGGAAATTTCCACTTCAAGACGTTCGATGGTCATTTCTTTCAGCTGCCAGACACGTGCAACTACGTTCTGGCGGT
GATGTGTGACGCCGCCATTTCTGATTTTAACATCCAGATGCAAAGAGACACTGTCAATGGCTCCATCTCATTTGGCT
CAGTCATAGTAAAGCTGGACGGGACAGTCATCAAAATCACAGACAGCGGCATCATGATGGACGATCAAGTGGTGTCT
GTTCCGATCAATCAGAAGGGAATTAAAATAGAAGGAAGCCCAACAAGTATTAAGATCTCCAGATATGGAATGACCAT
CTTTTGGGAAGAAGACAACTCTATAGCGATTGAACTTGCGGAGAAATATAAAGGGCTGACATGTGGACTCTGCGGAA
ACTACAATGGCGACAAGGATGATGATATGCCTGAGAGTGGTAAATGTCATGTTTTTGTAGTGTCATCCACATGGAAG
ATCTCAACACCAACTGAATCCTGTGAAGACGTTATACTTCCACCCAAGGATCAGTGTGACCAGAACACAATGGTTTG
CCAGCAGTACCTCAGCAGTCCAGGGTTTAGTGGCTGTTATGATGTGATGGACATGAAAATATTTGAAAAAGCCTGTG
TGAGCGACATGTGCCAATGTTACGGCAGCCATGACTGTCTGTGCAACACGCTAACAGAGATCTCACGACAATGCACA
CATGCTGGAGGACAACCAGGAACATGGAGGACAGAACAGCTCTGCCCAAAGATGTGTCCTATAAACCTGCAATACAT
GGAGTGTGGAGGTCCGTGTAAGAGCACTTGCTCTGATCCAACTGCACATCTAATGTGTAAAGATCACTGTGTGGATG
GCTGCTTCTGCCCTGAAGGGACCGTTGAAGACGACATTGGTCAGGGTGGTTGTGTTCCTGTAAACGAGTGTCCATGT
GTGCATGATGGCACAGTATATAAATCAGGAGAGTCTTACCAACAAGCTTGCAAAAAAATGTTTTTGTGCCGCTGGGCA
CTGGACCTGCACATACCTGGACTGTCCCGGCACATGCTCTGTTGTTGGAGGTTCTCACGTCACCACTTTTGATGGAA
AGAGCTTCACCTTCAGCGGGAACTGCGACTACATCCTCACTAAGCACTCTAACGACAGTGACTTTGCTGTTGTGGGA
AATCTGGCAAAATGTGAGCCAGCACGAACAGACACATGTCTTAATTCAGTTACCCTTGTCATCTCTGGAACCACTGT
AAGCTCTAGATTTTGTATCCTAACCACTCTTTTCATTGTCCCAGGTCCTGTGAGCATCTTCCAGCCCTCTTTGTCTT
TCATTATTGCTGATTTAAACAGTCTTCGTCTGGAAATCCAGCTGGCTCCAGTTATGCAGTTGTACATTGTAGCAAGC
ACTGAAGAGAAAGGAAAAATGACTGGACTGTGTGGAAACTACAATGATGTCCAATCAGATGACTTTAAAACAGATTT
GGGTATCACAGAGGGCACAGCGATATCTTTTGCCAACTTTTGGAAGAAAACCCCATATAGTTGTCCTGATCTTGAAA
ACACATTTGACAACCCCTGTAGCCTGAGTGTGGACACAGAGAAACTTGCTAAAGACTGGTGTTCCCGTCTGACTAAC
CAAAGTGGAGCCTTTTCTGCATGTCATTCAGAAATATGCCCAAAGATTTACTATGAAAGGTGCGTCTATGACACCTG
```

CAAGTGTGCAGATATCAGGAAGTGCATCTGTGCGGCTGTGTCCACTTATGCCCATGCTTGTGCTGCAAGAGGAATCA
TTCTCCAAGGCTGGATGGATTCAGACCCTTGTGAATCTGAGTGTTCAGAGAACATGAAGCACTCCTACGGTATGACT
GGCTGTGGCCTCACCTGCCGTTCTCTCAGTGGACAAGAAAACACCTGCCAAGGGTCCTTCACGCCTGTAGATGGCTG
TGTTTGTTCTGAGGGAACCTACCTTAATGAAGAAGGCATCTGTGTGCCTGCTGACCAGTGTCCTTGTTACAGCGGTG
ACCAGGTCACAAAACCTTCGGAAGTATCCCACGTAGATGGTCTAACATGCACCTGTAAACTTGGCAAACTGCACTGC
TCCAATCCTGAAAGTAAGTTTCAACCCTGTGTTGCTCCAATGGTGCTCTTCAAATGCTCAAATTATGAACCAGGAAA
AAAGGGCACAGAGTGTCAGAGAACATGCCAGAAACAAGACATTAACAATTGTGTGAGCACAGGGTGTGTATCAGGGT
GTATGTGTCCAGATAACCTTCTGGCTGATGGACAAGGTGGATGTGTGGAGAGGGAAAAATGTCCCTGTGTCCACAAT
GGAGTTACCTATTCATCTGGAGAACAAGTTCAAGAGGACTGTAACACCTGCACATGCAGAAATGGGATGTGGGATTG
CACGGAAAAAGAGTGTTACGGCACCTGCACCATCTATGGTGAAGGTCATTTCATGACATTTGATGGCAAGAAATATT
CTTTCCATGGAGACTGTGAACAAATCTTAGTCCATGATTACTGCAATACAGACCAGAGTCCATACTCTTTACGCCTT
GTCACAGAGAACATCCCATGTGGCACCTCTGGAACCATTTGCTCCAAGTCCATCAATCTTTTCTTTGGGAGATACAA
ACTGATTTTATCAGAGGAGAAAGAAATTCAAGTTGTTGAAAGCAATGGCACAGATTACCAATACCAAATCCATACTG
CTGGGATATACAATGTCATTGAGGTCAAAGGTCTGTTGAACTTAATCTGGGACAGTAAAACAAGTGTGATGCTCCAA
CTGCATCCCAAGTTTAAGGGCAAGGTGTGTGGATTGTGTGGAAACTTTGATGGTAATGCTAACAATGACTTTATGAA
ACATGATGGAGAAGAAGTCACCGATCCAGTGGTATTCGGGAATAGCTGGAAGTCGAATCCAGGCTGTCCGGATGTGA
GCAATGTTATGAACCCATGTGAAAAGAATCCACACCGGAGTGCCTGGGCTATCAAACAATGCAGCATCATTACAAGC
CCTGTCTTTTCAGACTGCCATTCACGTGTGGACTCTGGCCCATACTATGATGCCTGTGTAAAGGACACCTGTGCATG
CGATAGCGGGGGTGACTGTGATTGTTTCTGTACTGCAGTGGCTGCTTATGCTGCTGAATGCCGTAAAACCGGTGCTT
GCGTAGCATGGCGAAGCCCAAGTATTTGTCCTTTATTCTGTGACTATTACAACAACCCTCCAAGTGAATGTGAATGG
CACTATAAAACCTGCGGATCGACCTGCATGAAAACTTGCAGAAACCCATCAGGAACTTGCTCCAATCAAATTCCCCT
CCTAGAAGGATGTTTTCCTCAGTGTCCTTCTGAAAGACCATTTTTGAGAGAGGACAACAGGAAATGTGTTACAGAGG
CAGAATGTCTTTCTCTCTGTACTTATGATGGCAAGATATACACGACTGGAGATGTGATGTACGATACCACAGATGGG
AATGGTACATGCTTTACAGCTGTGTGTGGATCCAATGGAGAGATAATAAGAAGCATTAATAAATGTATGACAACAAC
TACACCATTCACGTTCACAACTCCACCCCAAACACCAACACCTTCATCACCCGAGACATATACTTCTGTTACTGCTG
CAACACCTTCGACAGAGAGGCCACCAACATCCGAATTGCAAAGGCCAAAAACAACTACTACAACTTCTGGTCCAACA
ACAAAA

muc5.3

F1 - IJ80 – cgagcaatatgcacagcact
R1 - IJ81 – gctcgagcagttgaaaaacc

TTTTCAGCAAGATTGTGACCAGGGGCAGTGGGATTGTTACATAAAGGATTGTCCCCATTCATGTTCAGTGGAGGGCG
GATCCCACATTACCACATTTGATGGAAAAGCATACACATTTCACGGAGACTGCACTTACGTCCTGTCCAAACATTGT
TCTGGCACAGAATTCACAGTTCTAGGAGACCTGGTGAAATGTGGTCTGACTGAGACCGAAACCTGTCTGAAGGCAAT
AACCGTGGCGTTGTTTGATGGAATTACTGTTATTAACGTGGACTCCAGCGGAAACGTTGAGGTCAACAGAATAGTTG
CTCAACTTCCACTTTTTACATCTTCAGTGAGTGTTTTCAAGCCCTCCTCATTCTACATCATCATTGAAGCTCCGCTG
ATTGGGCTTCGTCTGGAGATCCAGCTGGTGCCAATAATGCAGCTTTACATCACAGCCAATTCAGTGTATCAAGGAGA
AACATGTGGTCTTTGTGGCAATTACAACAACATGGAGGCTGATGACTTCACAACCATCGGTGGACTCCGAGAAGGAA
CAGCTCTGGATTTTGCCAACACCTGGAAAACAAGAGCAAGCTGCCCTGATGTTAAGAGAAGCTTTGAGAACCCGTGC
AGTCTGAGCACAGAAAACGAGCAATATGCACAGCACTGGTGTTCCCTGCTGTCTGATCCCTCGTCAGTTTTTGCTCC
GTGTCACACTGAAGTCAACCCTGAAGTCTACCAAGCAAACTGCATCTACGACAGCTGTAACAGTGAAAGGAGCCAGG
ACTGCAAGTGTGCGCATTATCATC

muc2.1

F1 – IJ125 – tgctgcttctggcgctttctggt
R1 - GY2 – cataaccaatttccccatcg

TCGGATGGCTCGAGTTTTTAGCTGCATGCTGCTTCTGGCGCTTTCTGGTATTCAGGTCGATTCAAAAAAGGTCTCTC
CTAGTAATCATGTCAACAGCATCTGCAGCATGTGGGGAAACTTCCACTTCAAGACCTTCGATGGGGACGTCTATCAA
TTCCCAGGGATGTGCGAGTACAACCTGGTGTCTGACTGCCAAAGCCTCATTCGACAGTTTTCGATTTACGTGAAGAG
GACTGAACGCAGCACTGGCCCAAAGATCAGCAGGGTCTCCATAACCATCAACGACATCGCCATTGAGTTGACCGAGA
ACCAAGTCAACGTCAATGAAGCAAAAGTGACGCTGCCCGTCCATGTGTCAGGAATCCTTGTGGAAGAAAACACAATT

```
TACACCAGACTCTATTCCAAAATGGGCATCACTGTCATGTGGAACAAAGATGATGCAGTCATGGTGGAACTGGACTC
AAAGTATTCCAATCGCACTTGTGGATTGTGTGGTGATTTCAATGGCATTCCTGTTTATAATGAGTTTATTCAATCAG
GACGAACTGTTGGATACACTGAGTTTGGGAATATGCACAGAGTTCCCAACCCAACTCATCAATGCGAGGATCCATTT
GAAAACGTGGATGAGCAAAATGTGGTGGACCAGTGTGAAAAATATCGGGCAGACTGCGCAGACCTTTTGGAGGATGA
AAAGTGGTCTTCTTGTAGCTGGGTTTTGGACCCTGAGGCTTATATTAAGGCATGTACAAATGATCTCCGCAACCGTC
AACCTGAAGACGAAGACACCACTGCACTCTGTGCAACTCTAACTGAGTATTCACGACAGTGTTCACACGCTGGAGGA
AATCCACCCGCATGGAGGACAGCAAAGTTCTGCAATGTGCAGTGCCCCTACAATATGGTACANNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGACTTGTGGCCTTTGTGGAAATTTCAACAAAGTGCTGGCTGATGA
CCTGAAGACCCCTCAGGGGGTGGTGGAGGGTACAGCGGTTTCCTTCGCAAACGCCTGGAAAGCCCAGTCCAACTGCC
CTGACCGTTCTGAGAGGATGGATGACCCCTGTTCCTACAGCAGTGATAGCGAGCACTTTGCTGAGCACTGGTGTTCC
AAGATGAAGGACAAGGAGAGCCTCTTTGCCCAATGTCACGCCACGGTCAATCCAGACAGCTACTACAAAAGATGCAA
ATACTCCAGCTGTACCTGTGAGAAGAGCGAGGACTGTCTGTGTACGGTTTTCTCCTCATATGCCCGGGCCTGTGCGG
CTAAAGGAGTAATCCTCGAAGGCTGGAGAGAAGTTGTGTGCGACAAGTATATAAAGGATTGCCCAGCCTCACAGAGC
TACTCCTATGAGCTGCAGAGCTGTCAGCGCACCTGTGTTTCTCTCGCCTCTGAGCGCCAGAGCTGCAATGTTGACTT
TGTTCCTGTTGACGGATGCGCATGCCCAGATGGACTGTACCAAGATGAGAATAATCTTTGTGTACCTATGGAAAAAT
GTCCATGTTACCACAACGGGCAGAAGATCCATCCTGGAAAGTCTCTCACTATTAGGGACGAACACTGTGTTTGCATG
AATGGAAAATTCCACTGTCGATCCTGGAAAACCCAAATTCTGGGATGTCCCTCTCCAAAGGTGTTCTTCAATTGCAG
TACAGCAGGCCCAGAAGAATATGGCTTGGAATGTGCAAAGACTTGCTCGCATCAAGAAGTTGACTGTTTCTCTCTAG
ACTGTCAGTCTGGCTGCCAGTGTCCTTTTGGTCTACTGGATGATGGCAGAGGACACTGCGTTCAACCGTACAACTGC
CCATGTAAACATGATGGACAGTATTATGCACCAGGATCTGTGATTGTTAAAGATTGTAATAAATGCATTTGTCGATG
GGGAAATTGGTTATG
```

muc2.2

F1 - IJ111 – actgtccgtgtcctcatggt
R1 - IJ112 – tctcctccagtgtcacatgc

```
ACTGTCCGTGTCCTCATGGTGGACAACTTTACCAACCAGAGACTGAGATAACTGTAGACTGCAATAAGTGCACTTGT
CAATCTGGAATATGGACATGTACAAAAAATAAATGTCCAGGAACCTGCACTATATATGGAAGCGGACATTATAAAAC
TTTCGACAATCAAAGGTTTGGATTCCAAGGATCCTGCAGTTATATAGCAGCTCAGGATATGCGTGGAAATAGGACAG
GAATATTTCAAGTAATAACTGAAAATATTCCTTGTGGAAGAACAGGGACCACATGCTCAAAGTCTGTCTTTATTCAA
TTTGGGAGAGTGTCCCTACACCTTTCAGATGAGCTTGAAACACAATATTTAATGGAAAGTGGCGCACTAATTAGAAA
AATTGGAATGTACCTTGTCATTGATACTAACATTGGGCTGACAGTTTTGTGGGACCGCAAAACCACCATTCACATCA
TTCTGCAACCTGAGTACATGGGAGAAGTATGCGGCCTATGTGGAAATTTTAACGGAGATGGAAAAGACGACTTCATC
ACTAAAGGTGGTGTGCAAACTTCAAATATAATTGAGTTTGTTGACAGCTGGAAGCAGAAAACCTCCTGCCCTGATGC
TGCACCAGACTTTGACCCATGCGTTACGAACCTCTATAGAGAGGAATGGGCAGTATTAAAATGTAGCATCATAACGA
GTGATACATTCAAAGACTGCCAAAAAAAGGTGGATCCAACACCGTATTATGACAACTGTTTGAAGGACACGTGTGCA
TGTGACACTGGAGGAGA
```

F2 - IJ113 – accacaaccaaacccatgtt
R2 - IJ114 – cctctcgaatgctggatctc

```
ACCACAACCAAACCCATGTTATGAATGTATAAACATAAATGGAAGTTTCATCCCTGTTCAACGGGATTGTCTTCCAC
CGCAAAACATCACCTGTGAAAATCACCAGCCAAGCGTGCTGCTAGAGGATAAGCACGAATGCTGCCCTTATTACGTC
TGTGATTGCTTTTGCCAAGGCTGGGGTGACTCCCATTACATCACATTTGATGGACTCTATTACAATTACCAAGGAAA
```

CTGCACTTATATTTTAATGAAGGAAATTAGCCCTCAATTCAATTTAGCAATCTACATTGAGAATGTCTACTGTGACC
CTGCTGAGAATGTATCCTGTCCCAGATCCCTTATTGTGTCTTACAACAAACAAAACATTACACTTAGAAGTCACAGG
CTAATGGGAGGAGCAAATCTAGAGGCTCTAGAGGGAAATATTATGCTAGCACTGCCGTATGATAATAATGGCCTGAG
GGTTATTAGTTCAGACCTAAACTTATTACTAGAGATTCCAAAACTCAACGCCTATGTACTATTTGGAGTCTCTGGCT
TCAGCATCAACCTGCCATTCCAGTATTTTGGAAAAAACACAGAAGGCCAATGTGGAACATGCAACAACAACCAGAAT
GATGATTGCCCTGGATCCTTGGCAAACGACTGTAAAATAATGGCAGAGAACTGGAATACTACAGTTTGCACTCCACC
AGCTCCATTACCTACTCCTAACTCACCCCATGTCCACCCTGACTGCTACCTACTAAACACCACGTTTGAAGCGTGCC
ACTCCCATGTGCCTCCTGAAAACTTCTTTTCAGCCTGTCAATATGACAGCAGTAATAACCCTGCTGAAGTGTGTGCC
AGTCTGCAGAGCTACGCAAGCGTCTGTGCAATGTTTGGGGTCTGCATATACTGGAGGAACTACACAAGCCAATGCAA
TATTGGTTGTTCAGAAGATAAAGTGTTCTTGCCTTGTGGTTCATTTGAACCGCCAACATGTAGAGATAGACAAATGG
AGAGCAACATTACAGTGTCCACAGAAGGCTGCTTCTGTAGTCAGAACACAACACTCTTCACCAAGGATTCTGGAGTC
TGTGTGCCACAACCCTGCGGATGCTTGGACTCATCTGGGAAACCACGTAGGTTTAATGAGAATTTTACGCACAACTG
TGAGGACTGTGTCTGTGATAAGGCCAGGATGTCAGTCATCTGTGAACCCAAGAAGTGTCCCGATGTGCCCCCTGTGA
ACTGCACTGAGCCCTTCATGGTGGTCAATGTCACCGATCCATCAGAGCCATGCTGCAGCAGACAAGTTTGCAATTGT
AGCATGTGCCCATTTTCGGAGGAAAAGTGTGAGGTTGGATATGAACTGATTGTGCACCCTTACGGAGACTTCTGTCC
AGAAATCAGATGCGAGTCAAAAGAAGTGTGTGCCTACAACAACACAGAATATAAGCCTGGGAGTAAAATCCACGTAG
GGCCTTGTGTTAAATGTAACTGTGAATGGGACCCATGGACTGAACTATATCAAATGAATTGTTCCTATGAAGTTTGC
CCTGATGATGAGTGTTATGAGGGGTTTGAATATGTGAAACAAGATGGAGAGTGCTGTGGAACATGTAAACTAATGAG
CTGCATTTATTATAGAGAGAACAACACTAAACACATTCTTCATGATGGAGATGTCAAGAATTTTACATGTGAAACTG
TTTCCTGTCATGAAATAAATGGCTCATTTTTGATTGAAAAAACCAAAACAGTGTGTTCGTACTCGAGTGTGCACGTC
TGCGAGTATGGGTTTGAATATGTGATGAAAGAAGGCGAGTGCTGTGGAACATGTGAACCAGTGGCCTGTATTTATCA
ATCACCAAACAACACCATACATATTCTCAAGGATGGACAAGAGTTTAAGTATAAATGTGAAACTGTTACCTGCAATC
ATATTAATGACTCATTTGAGATGAACATAACCAACATGGAATGTATTCAATTGAGTTCAGAGGACTGTGCGTCTGGC
TTTCAATATATGAAAAAAGATGGAGAGTGCTGTGGGTCGTGTGTACAAGTGGCCTGTATTTATGATGCTCCAGACAA
AACCACACATGTTCTCCAGGAAGGACTTTCATACAATATATCATGCATGAATGTCACATGTCTGAAAAGGAATGACG
TGTTCACGATAAAGGAAGAAATGAAAAAATGCCCTTCCTTTAATCAAGATGACTGTGTGGAGGGAACAAAGCAATTT
GATGAGGATGGATGTTGCCCAAACTGTGAAACTCGTAACTGTGCTCCTGTGAAGAACGTCACTCGTCTGAAAGTCAA
TAATTGCACTTCAACTGAAGATGTTGAAATTACATCCTGCGCTGGTCACTGTGACTCCAAATCAACTCTGGATCTAC
ATGGATACATGATGAACAGCTGTTCCTGCTGCCAAGCAGAGAGATCCAGCATTCGAGAGG