

# CLUSTAL 2.1 MULTIPLE SEQUENCE ALIGNMENT

File: D:/New\_LifeScience/collaborations/WangHongyan/GATA1 to GATA6/GATA345 Date: Fri Mar 23 22:31:26 2012  
 Page 1 of 3

GATA-4 Human  
 ENSPTRG00000019993 Chimpanzee  
 ENSBTAG00000005425 Cow  
 ENSRNOG00000010708 Rat  
 ENSMUSG00000021944 Mouse  
 ENSGALG00000016662 Chicken  
 ENSXETG0000000673 Frog  
 ENSTRUG00000012112 J Puffer  
 ENSTNIG00000016369 T Puffer  
 ENSDARG00000035759 Zebrafish  
 GATA-5 Human  
 ENSPTRG00000013711 Chimpanzee  
 ENSCAF00000012709 Dog  
 ENSMUSG00000015627 Mouse  
 ENSRNOG00000036732 Rat  
 ENSGALG00000005352 Chicken  
 ENSTNIG00000002769 T Puffer  
 ENSDARG00000017821 Zebrafish  
 GATA-6 Human  
 ENSBTAG00000005734 Cow  
 ENSMUSG00000005836 Mouse  
 ENSRNOG00000023433 Rat  
 ENSTRUG00000012674 J Puffer  
 ENSTNIG00000014837 T Puffer  
 ENSDARG00000035831 Zebrafish  
 ENSXETG00000003144 Frog  
 FBgn0003117\_Fly

```

-----
---MALT DGGWCLPKRFGAAG-ADASD-SRAFPAREPSTPPPISSSSSS--CSRGGERGPGGASNCITPOLDTEAAGPPARSLLLSSYASHPFPGAPHGFSAPGVAGPGGNLSSWEDLLLFIDLDCAATASKLLWSSR--GAKLSP 138
---MALTDGGWCLPKRLGAAPPADASD-SGAFPAREPSTPPPISSSSSSC--CSRGGERGPGGASNCITPOLDTEAAGPPARSLLLSSYASHPFPGAPHGFSAPGISGPGALSSWEDLLLFADLDCAATASKLLWSSR--GAKLSP 140
---MALTDGGWCLPKRFGAAA-ADAGD-SGFPFAPREPSSPLPISSSSSS--CSRGGDRGPCGASNCITPOLDTEAVAGPPGRSLLLSPYASHPFAAAHGAAPGVAGPGSALSTWEDLLLFIDLDCAATASKLLWSSR--GAKLSP 138
---MALTDGGWCLPKRFGAAS-ADAGD-SGFPFAPREPSSPLPISSSSSS--CSRGGDRGPCGASNCITPOLDTEAVAGPPGRSLLLSPYASHPFAAAHGAAPGVAGPGSALSTWEDLLLFIDLDCAATASKLLWSSR--GAKLSP 138
---LLPTLTKGKCSWMDLGENSWSMVKREVS--PGSPAEQI--VLPGDSRRDGPISDDLRTF-PSLDLALGHRPDPGRSLHSYVHFCHHN-----NTLTTVEDIPLFTDLDGOG--KLVLPGG--AHKAG 115
-----SWMDLGENSWSMVKREVS--PGSPAEQI--VLPGDSRRDGPISDDLRTF-PSLDLALGHRPDPGRSLHSYVHFCHHN-----NTLTTVEDIPLFTDLDGOG--KLVLPGG--AHKAG 105
---MDLGDNSWSMVKREVS--PGSPAEQI--VLTNDRR--DRLRSAPATHLDPVGSRRRABGRPLHSYVHFCHPN-----NALPSSDDVTLFTDLDGOGN--KLIISGGHPRETHKGG 104
MSWERHPOESRLRIPISORPHITWMDLNEDVSGGTRLSANSPFPCPIKQSPGTPPPSPFCQLGSS--ECLRLPDRGPHITIS--AEGRPAPPEYHPLG-----LSHPEDLILFRDLHDIN--KLIVPNR--ETRYSS 125
-----
1.....10.....20.....30.....40.....50.....60.....70.....80.....90.....100.....110.....120.....130.....140.....150
    
```

GATA-4 Human  
 ENSPTRG00000019993 Chimpanzee  
 ENSBTAG00000005425 Cow  
 ENSRNOG00000010708 Rat  
 ENSMUSG00000021944 Mouse  
 ENSGALG00000016662 Chicken  
 ENSXETG0000000673 Frog  
 ENSTRUG00000012112 J Puffer  
 ENSTNIG00000016369 T Puffer  
 ENSDARG00000035759 Zebrafish  
 GATA-5 Human  
 ENSPTRG00000013711 Chimpanzee  
 ENSCAF00000012709 Dog  
 ENSMUSG00000015627 Mouse  
 ENSRNOG00000036732 Rat  
 ENSGALG00000005352 Chicken  
 ENSTNIG00000002769 T Puffer  
 ENSDARG00000017821 Zebrafish  
 GATA-6 Human  
 ENSBTAG00000005734 Cow  
 ENSMUSG00000005836 Mouse  
 ENSRNOG00000023433 Rat  
 ENSTRUG00000012674 J Puffer  
 ENSTNIG00000014837 T Puffer  
 ENSDARG00000035831 Zebrafish  
 ENSXETG00000003144 Frog  
 FBgn0003117\_Fly

```

-----
---MYQSLAMAANHGPPPGAYEAGGPGAFMHGAGA---ASSPVYVTPTR-VPSVVLGLSY-LOGGGAGSAGGGASGGSSGGAASGAGPGTIQGGSPGWSQA-GADGAAYTP-----PPVSP----- 106
---MYQSLAMAANHGPPPGAYEAGGPGAFMHGAGA---ASSPVYVTPTR-VPFVVLGLSY-LOGGGAGSAGGGASGGSSGGAASGAGPGTIQGGSPGWSQA-GADGAAYTP-----PPVSP----- 106
---MYQSLAMAANHGPPPGAYEAGGPGAFMHGAGA---ASSPVYVTPTR-VPSVVLGLSY-LOGGGGGAASGAASGGSSGAAPSGAGPGTIQGGSPGWSQA-GAEGAAYTP-----PPVSP----- 106
---MYQSLAMAANHGPPPGAYEAGGPGAFMHGAGA---ASSPVYVTPTR-VPSVVLGLSY-LOGGGGAAAGATSGGSSGAGPSGAGPGTIQGGSPGWSQA-GAEGAAYTP-----PPVSP----- 106
---MYQSLAMAANHGPP--SYEGAG--GFMHGAAA---AS-PVYVTPTR-VPSMLPISLPI-LPSSGSSQQASPVSSHS-----IWTQP-GAESAAANPG-----SS-HPPVSP----- 88
---MYQSIAMAANHGPP--AYEGAG--GYMHSATA---ATSPVYVTPTR-VSMIPISLPI-LPSSGSSQQASPVSSHS-----MWAQA-GAEPANPG-----SS-HPPVSP----- 89
---MYQGI-MSNHGGPP--SYEAG--FLHNTSA---ATSPVYVTPTRVVTMPALPI-LQTFGASQQSSPVSSHS-----AWAQP-GSDSVSNYS-----HSPVSP----- 84
---MYQGI-MSSHGPP--SYEAG--FLHNTSA---ATSPVYVTPTRVVTMPALPI-LQTFGASQQSSPVSSHS-----AWAQP-GSEPVSNYS-----HSPVSP----- 84
---MYQGVMAANHGAA--SYEG--FLHNS---TSPVYVTPTR-VTPMLQALPI-LQAP--QQSSPASGHS-----GWAQP-GVETISVNS--FGHHSFVS----- 82
---MYQSLALAASPRQAAY--ADSGS--FLHAPG---AGSPMFVPPAR-VP--SMLS-LSGCEPSPQPELA-----ARPGWAQTATADSAFG-----PGSPHPAAHPPGA 90
---MYQSLALAASPRQAAY--ADSGS--FLHAPG---AGSPMFVPPAR-VP--SMLS-LSGCEPSPQPELA-----ARPGWAQTATADSAFG-----PGSPHPAAHPPGA 90
---MYPSLALAPSPGQAAY--TDSGA--FLPAPA---AGSPVFPVPPAR-VP--PMLPI-LPCEPSPQAPGRG-----AHPGWAQ--AADSAFG-----PGSPHPAAHPPGA 88
---MYQSLALAASPRQAAY--ADSGA--FLHSSG---AGSPVFPVPPAR-VP--SMLPI-LPCEPSPQAPALA-----AHSWTQVAADSAFG-----PGSPHPAAHPPGA 90
---MYQSLALAASPRQAAY--ADSGA--FLHSSG---AGSPVFPVPPAR-VP--SMLPI-LPCEPSPQAPALA-----AHSWTQVAADSAFG-----PGSPHPAAHPPGA 90
---MYQGLALAFNHGQASVHSDSGN--FLHSS---AGSPVYVTPTR-VPSVLQTLPI-LQSCPEP-HQSHLG-----HPPGWAQS-SGETAFN-----AGSPHPSPG 85
---MYQSLALSFN--QPPPHDTGT--YLHPP---ASSPVYVTPSR-VPAVLPTLPI-LQCEAPPQSHGLGGHHGLGGHHC-----LGGHHGPPFA-SADGPPFP-----FSSPLPHG 97
---MYSLALSFN--PSPVHSDSGN--YIHPS---ASSPVYVTPTR-VPAMLQTLPI-LQTCESSHQAHTIS-----SHHAWPQT-GTDNSSFN-----PGSPHPPPG 85
FAAEQP-EMYQTLAALSSQGAAY-DGAPGG-FVHSAAAAAAAAAAASSPVYVTPTR-VGSMPLGPIVHLGSSGSGPANHAGG-----AGAHPGWQA-SADSPYGGGGAGGGGAGPGGAGSAAHVVSAR 262
FAAEQP-EMYQTLAALSSQGAAY-DGAPGG-FVHSAAAAAAAAAAASSPVYVTPTR-VGSMPLGPIVHLGSSGSGPANHAGG-----AGAHPGWQA-SADSPYGGGGAGGGGAGPGGAGSAAHVVSAR 262
FAAEQP-EMYQTLAALSSQGAAY-DGAPGG-FVHSAAAAAAAAAAASSPVYVTPTR-VGSMPLGPIVHLGSSGSGPANHAGG-----AGAHPGWQA-SADSPYGGGGAGGGGAGPGGAGSAAHVVSAR 260
FAAEQP-EMYQTLAALSSQGAAY-DGAPGG-FVHSAAAAAAAAAAASSPVYVTPTR-VGSMPLGPIVHLGSSGSGPANHAGG-----AGAHPGWQA-SADSPYGGGGAGGGGAGPGGAGSAAHVVSAR 258
LLVDP--TDMYQTLAIAAASQSGY-DSSGG--YMHNSN-----NSPVYVTPSR-VGSMIPISLPI-LQAGG-AQPNHA-----VSSHVWSQS-TTESPSVS 209
LLVDP--TDMYQTLAIAAASQSGY-DSSGG--YMHNSN-----NSPVYVTPSR-VGSMIPISLPI-LQAGG-AQPNHA-----VSSHVWSQS-TTESPSVS 209
LIVDP--TDMYQTLAIAAASQSGVNDTPSAG-YMHNSN-----TSPVYVTPSR-VGSMIPISLPI-LQAGG-AQPNHA-----VSSHVWSQS-TTESPSVS 198
ALPEPPAGDMYQTLTITAAQGLGV-DHSQGT-FMHSAA-----SSPVYVTPSR-VGSMIPISLPI-LQAGG-AQPNHA-----VSSHVWSQS-TTESPSVS 199
ALPEPPAGDMYQTLTITAAQGLGV-DHSQGT-FMHSAA-----SSPVYVTPSR-VGSMIPISLPI-LQAGG-AQPNHA-----VSSHVWSQS-TTESPSVS 219
MGILLSDDGDSIDQOSTRDYPHFSGDYQN--VLLAASASTASASATHVAAVKMYSSAVAAVYTLAAAG-----AASAGVGVGSG----- 82
-----
160.....170.....180.....190.....200.....210.....220.....230.....240.....250.....260.....270.....280.....290.....300
    
```



CLUSTAL 2.1 MULTIPLE SEQUENCE ALIGNMENT

File: D:/New\_LifeScience/collaborations/WangHongyan/GATA1 to GATA6/GATA345.Date: Fri Mar 23 22:31:26 2012

Table of multiple sequence alignment for GATA-4 across various species including Human, Chimpanzee, Cow, Rat, Mouse, Chicken, Frog, Puffer, and Zebrafish. It shows amino acid sequences with conserved regions highlighted and a scale at the bottom from 310 to 450.



A second table of multiple sequence alignment, likely for GATA6, showing sequences for the same set of species as the first table. It includes a scale at the bottom from 460 to 600 and uses different highlighting to indicate conserved regions.



# CLUSTAL 2.1 MULTIPLE SEQUENCE ALIGNMENT

File: D:/New\_LifeScience/collaborations/WangHongyan/GATA1 to GATA6/GATA345 Date: Fri Mar 23 22:31:26 2012  
 Page 3 of 3

GATA-4 Human	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	442
ENSPTRG00000019993 Chimpanzee	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	399
ENSBTAG00000005425 Cow	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	442
ENSRNOG00000010708 Rat	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	441
ENSMUSG00000021944 Mouse	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	442
ENSGALG00000016662 Chicken	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	411
ENSXETG00000000673 Frog	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	394
ENSTRUG00000012112 J Puffer	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	386
ENSTNIG00000016369 T Puffer	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	386
ENSDARG00000035759 Zebrafish	EMRPIKTEPGLSSHYGHSVSSVOT-FSVSAMSGHGPSIHPVLALKLSPQGYASPVVSQSPQSSKODSWNSLVLADSHGDIITA-----	352
GATA-5 Human	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	397
ENSPTRG00000013711 Chimpanzee	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	397
ENSCAFG00000012709 Dog	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	400
ENSMUSG00000015627 Mouse	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	404
ENSRNOG00000036732 Rat	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	404
ENSGALG00000005352 Chicken	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	391
ENSTNIG00000002769 T Puffer	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	392
ENSDARG00000017821 Zebrafish	SLASPVCPGPMMA--PQASGQEDDELAPCHLEFKFEPEDFAFPSTAPSPQAGL-----RGALRQEAWCALALA-----	383
GATA-6 Human	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	595
ENSBTAG00000005734 Cow	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	499
ENSMUSG00000005836 Mouse	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	589
ENSRNOG00000023433 Rat	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	587
ENSTRUG00000012674 J Puffer	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	511
ENSTNIG00000014837 T Puffer	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	497
ENSDARG00000035831 Zebrafish	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	501
ENSXETG00000003144 Frog	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	524
FBgn0003117_Fly	NTSPPTQPTASGAGAPVMTGAGE-SINPENSELKYSQDGLYIGVSLSPAEEV-----TSSVRQDSWCALALA-----	467



GATA-4 Human	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	442
ENSPTRG00000019993 Chimpanzee	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	399
ENSBTAG00000005425 Cow	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	442
ENSRNOG00000010708 Rat	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	441
ENSMUSG00000021944 Mouse	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	442
ENSGALG00000016662 Chicken	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	411
ENSXETG00000000673 Frog	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	394
ENSTRUG00000012112 J Puffer	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	386
ENSTNIG00000016369 T Puffer	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	386
ENSDARG00000035759 Zebrafish	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	352
GATA-5 Human	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	397
ENSPTRG00000013711 Chimpanzee	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	397
ENSCAFG00000012709 Dog	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	400
ENSMUSG00000015627 Mouse	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	404
ENSRNOG00000036732 Rat	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	404
ENSGALG00000005352 Chicken	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	391
ENSTNIG00000002769 T Puffer	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	392
ENSDARG00000017821 Zebrafish	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	383
GATA-6 Human	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	595
ENSBTAG00000005734 Cow	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	499
ENSMUSG00000005836 Mouse	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	589
ENSRNOG00000023433 Rat	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	587
ENSTRUG00000012674 J Puffer	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	511
ENSTNIG00000014837 T Puffer	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	497
ENSDARG00000035831 Zebrafish	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	501
ENSXETG00000003144 Frog	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	524
FBgn0003117_Fly	.....610.....620.....630.....640.....650.....660.....670.....680.....690.....700.....710.....720.....730.....740.....750	540

HFHAAA...VAAVGVKTEASATNYDVNNCYFGGTFGALGGAATTTAMAGGAASELAGYHHCHNVIQAAKLMATSS  
 .....760.....770.....780.....790.....800.....810.....820...