**Elucidating the Role of *MLL1*** **nsSNPs: Structural and Functional Alterations and Their Contribution to Hematological Malignancies Susceptibility**

**Hakeemah H Al-nakhle 1, \*, Hind S Yagoub 1,2, Rahaf Y Alrehaili 1, Ola A Shaqroon 1, Minna K Khan 1, Ghaidaa S Alsharif 1**

**Figure S1**. Domain identification of *MLL1* protein using InterPRO server. IPR016569 indicates the *MLL1* protein (1-3969 aa), the Post SET domain (3953-3969); SET domain (3829-3951); FY rich C (3666-3753); Znf PHD4 finger (1932-1978); Bromodomain (1633-17767); Znf PHD3 finger (1568-1625); Znf PHD2 finger (1481-1531); Znf PHD1 finger (1433-1480) and Zinc finger, CXXC (1147-1195).

|  |
| --- |
| 1          11         21         31         41         **MAHSCRWRFP ARPGTTGGGG GGGRRGLGGA PRQRVPALLL PPGPPVGGGG****eeeebebebe eeeeeeeeee eeeeebeeee eeeebeebbb beeeeeeeee**  **ff**     **f f** **ff**                     **f**  **f**               |
| 51         61         71         81         91         **PGAPPSPPAV AAAAAAAGSS GAGVPGGAAA ASAASSSSAS SSSSSSSSAS****eeeeeeeebb bbbbbeeeee eeeeeebbbb bbbbeeeeee eeeeeeeeee**                                                       |
| 101        111        121        131        141        **SGPALLRVGP GFDAALQVSA AIGTNLRRFR AVFGESGGGG GSGEDEQFLG****ebebbbebbe ebebbbebbb bbbebbeebe ebbbebeeee eeeeeeebbe**             **fsss** **s**           **s**                    **s** **f** |
| 151        161        171        181        191        **FGSDEEVRVR SPTRSPSVKT SPRKPRGRPR SGSDRNSAIL SDPSVFSPLN****bbeeeebebe eeeeeeeeee eeeeeeeeee eeeeebeebb eeeeebeeee**  **ff**                    **ffffffff**                      |
| 201        211        221        231        241        **KSETKSGDKI KKKDSKSIEK KRGRPPTFPG VKIKITHGKD ISELPKGNKE****eeeeeeeeeb eeeeeeeeee eeeeeeebee bebebeeeee eeeeeeeeee**        **f**          **ff** **fffff**      **f**      **f**       **f**     |
| 251        261        271        281        291        **DSLKKIKRTP SATFQQATKI KKLRAGKLSP LKSKFKTGKL QIGRKGVQIV****eebeebeeee eeebeebeeb eebeeeeeee beeeeeeeeb ebeeebbebb**   **f**   **f** **f f**   **f**   **fs ff** **f**  **ffff** **ff**  **f**  **f**  **f** **f**        |
| 301        311        321        331        341        **RRRGRPPSTE RIKTPSGLLI NSELEKPQKV RKDKEGTPPL TKEDKTVVRQ****eeeeeeeeee ebeebeebbb eeeeeeeeee eeeeeeeeee eeeeeeebee****ffffffff** **f f** **f**             **f** **f**   **f**              **f**   **ff** |
| 351        361        371        381        391        **SPRRIKPVRI IPSSKRTDAT IAKQLLQRAK KGAQKKIEKE AAQLQGRKVK****eeeebeebeb beeeeeeebe bbeebbeebe eeeeeebeee bbeeeeeeee****ffff** **ffsf**   **f**  **ffff**   **ssffssf** **s**  **ffffff** **fff ssffff** **f**   |
| 401        411        421        431        441        **TQVKNIRQFI MPVVSAISSR IIKTPRRFIE DEDYDPPIKI ARLESTPNSR****eebeebeebb bbbbbbbbbe bbebeeebbe eeeeeeebeb bebeeeeebe****ffsffsffss ssssss** **ssf ssfsfffssf ffff** **ff** **f**   **f** **f** **f**   **f** |
| 451        461        471        481        491        **FSAPSCGSSE KSSAASQHSS QMSSDSSRSS SPSVDTSTDS QASEEIQVLP****bebbebeeee eeeebeeebe ebeeeeeeee eeeeeebeee ebbeebeebe**    **fs** **fff ffff** **fffsf f** **ffffffff fff** **ffs** **ff fssf**       |
| 501        511        521        531        541        **EERSDTPEVH PPLPISQSPE NESNDRRSRR YSVSERSFGS RTTKKLSTLQ****eeeeeeeeee eebeeeeeee eeeeeeeeee bebbeeeeee ebbeebeeee**                              **ff** **f**            **ff**      |
| 551        561        571        581        591        **SAPQQQTSSS PPPPLLTPPP PLQPASSISD HTPWLMPPTI PLASPFLPAS****eeeeeeeeee eeeeeeeeee eeeeeeebee ebebbbbeeb ebbbebbebb**       **f** **f**  **fff**   **f**    **f**        **f**               **s**     |
| 601        611        621        631        641        **TAPMQGKRKS ILREPTFRWT SLKHSRSEPQ YFSSAKYAKE GLIRKPIFDN****eeebeeeeee bbeeeebebb bbeebeeeee bbebbebbee ebbeeebbee**      **ff** **f ssffffsfss s** **f**  **f**   **f** **sfssfssff fssfff** **sff** |
| 651        661        671        681        691        **FRPPPLTPED VGFASGFSAS GTAASARLFS PLHSGTRFDM HKRSPLLRAP****beeeeeeeee bbbbbbbeee eeebebebbb bbeebbebeb eeebebbebe****sffffff** **ff sss**                **s**    **f**   **fs**   **fffsf** **sfsf** |
| 701        711        721        731        741        **RFTPSEAHSR IFESVTLPSN RTSAGTSSSG VSNRKRKRKV FSPIRSEPRS****ebeeeebbbe bbebbbbeee eeeeeeebee eeeeeeeeeb eeebeeeeee****fsffffsssf s** **fsss**     **f**             **f** **f** **f**       **f**   **ff** |
| 751        761        771        781        791        **PSHSMRTRSG RLSSSELSPL TPPSSVSSSL SISVSPLATS ALNPTFTFPS****eeeebeeeee ebeeeebeeb eeeeeebebb ebbbbebbbe bbebebebee****ffffsfffff f**                  **s**           **f ssf**    **s**   |
| 801        811        821        831        841        **HSLTQSGESA EKNQRPRKQT SAPAEPFSSS SPTPLFPWFT PGSQTERGRN****eebeeeeeee eeeeeeeeee eeeeeeeeee eeeebbbbbe eeeeeeeeee**                                     **ss**                |
| 851        861        871        881        891        **KDKAPEELSK DRDADKSVEK DKSRERDRER EKENKRESRK EKRKKGSEIQ****eeebeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee**                   **ff**       **f** **f fffff**  **f** **f**  **f** **f**      |
| 901        911        921        931        941        **SSSALYPVGR VSKEKVVGED VATSSSAKKA TGRKKSSSHD SGTDITSVTL****ebebbbebee eeeeeeeeee beeeeebeee eeeeeeeeee ebeebbbbbb**                              **f**   **ff** **f**                 |
| 951        961        971        981        991        **GDTTAVKTKI LIKKGRGNLE KTNLDLGPTA PSLEKEKTLC LSTPSSSTVK****eeebebeeee beeeeeeeee eeeeeeeeeb eeeeeeeeee beeeeeeebe**      **f** **f**    **ff** **f**                                      |
| 1001       1011       1021       1031       1041       **HSTSSIGSML AQADKLPMTD KRVASLLKKA KAQLCKIEKS KSLKQTDQPK****ebbeebbbbb bebeebeeee eebbebbeeb eeebbebeee eeeeeeeeee**    **f**       **f** **ff**    **f ffss** **ss** **fs fffs** **fsff**  **ffff**    **f**  |
| 1051       1061       1071       1081       1091       **AQGQESDSSE TSVRGPRIKH VCRRAAVALG RKRAVFPDDM PTLSALPWEE****eeeeeeeeee beeebeebee bbebbbbebe eeeebbeeeb eebeebeeee** **f** **fffffff** **fffsffsff ssfssssfsf ffffssfffs ffsffsffff**26 |
| 1101       1111       1121       1131       1141       **REKILSSMGN DDKSSIAGSE DAEPLAPPIK PIKPVTRNKA PQEPPVKKGR****eeebbeeeee eeebebeebe ebeeeeeeee ebeeeeeeee eeeeeeeeee****fff** **sfffff fffsf** **ffsf** **sff** **fff** **f f** **ff** **ff** **f**   **fffffffff** |
| 1151       1161       1171       1181       1191       **RSRRCGQCPG CQVPEDCGVC TNCLDKPKFG GRNIKKQCCK MRKCQNLQWM****eeeeeeebee eeeeeeeeee eeeeeeeeee eeebeeebbe eeebeebeeb****fffffffs** **f ffffffff** **f ffffffffff fffsfffssf** **ff** **ffsffs** |
| 1201       1211       1221       1231       1241       **PSKAYLQKQA KAVKKKEKKS KTSEKKDSKE SSVVKNVVDS SQKPTPSARE****ebeeebeeeb eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee****fsf**  **sfff**  **f**  **f** **fff**   **f**   **f**                   **f**        |
| 1251       1261       1271       1281       1291       **DPAPKKSSSE PPPRKPVEEK SEEGNVSAPG PESKQATTPA SRKSSKQVSQ****eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee**           **f**                                     **ff**    |
| 1301       1311       1321       1331       1341       **PALVIPPQPP TTGPPRKEVP KTTPSEPKKK QPPPPESGPE QSKQKKVAPR****eeeeeeeeee eeeeeeeebe eeeeeeeeee eeeeeeeeee eeeeeeeeee**                 **f**         **f** **fff**                      |
| 1351       1361       1371       1381       1391       **PSIPVKQKPK EKEKPPPVNK QENAGTLNIL STLSNGNSSK QKIPADGVHR****eeeeeeeeee eeeeeeeeee eeeeeebebb bebeeeeeee eeeeeeebbb**         **f**   **f**        **f**       **s**    **f**      **f**    **ffsss** |
| 1401       1411       1421       1431       1441       **IRVDFKEDCE AENVWEMGGL GILTSVPITP RVVCFLCASS GHVEFVYCQV****bebebeeebe beebbebbbb bbbbbbbbbb ebbbbbbbee eebebbbbeb****sfsfsf** **f**       **s**  **s**      **ss** **s** **s**  **fsss** **sssff f**  **fss** **sfs** |
| 1451       1461       1471       1481       1491       **CCEPFHKFCL EENERPLEDQ LENWCCRRCK FCHVCGRQHQ ATKQLLECNK****ebeebeebbb eeeeeebeee beebbbeebe bbebbeeeee ebeebeebee****fsffsf** **sss ff** **fff** **f** **f** **ffsssffsf ssf** **sffff**    **ff** **ffs** **f** |
| 1501       1511       1521       1531       1541       **CRNSYHPECL GPNYPTKPTK KKKVWICTKC VRCKSCGSTT PGKGWDAQWS****beeeeeeebe eeeeeeeeee eeebebbbee bebeeeeeee eeeeeebebe** **fffffffsf fff** **ffffff ff**  **fs** **sff s** **sffff** **ff fff**  **fsf** **f** |
| 1551       1561       1571       1581       1591       **HDFSLCHDCA KLFAKGNFCP LCDKCYDDDD YESKMMQCGK CDRWVHSKCE****eebebbeebb ebbeeeebbe ebeebeeeee eebebeebee beeebebebe****ffsf** **sffss fs** **ffff** **sf fsffsfffff ffsfsffs** **f** **f** **f** **f** **fsf** |
| 1601       1611       1621       1631       1641       **NLSGTEDEMY EILSNLPESV AYTCVNCTER HPAEWRLALE KELQISLKQV****eeeeeeeebb ebbeebeeeb bbbbbebeee eeeebebebe eebeeebeeb** **ff**    **fss f** **sf** **sfff**  **ssss** **fsf**   **ffffsf** **fs**   **f** **f** **f**     |
| 1651       1661       1671       1681       1691       **LTALLNSRTT SHLLRYRQAA KPPDLNPETE ESIPSRSSPE GPDPPVLTEV****beebbeeebb eebbebeeee eeeeeeeeee eebeeeeeee eeeeebbeeb** **ffssfffs**  **ffssf** **ff**   **ffffffffff ff** **fffffff fffff** **sffs** |
| 1701       1711       1721       1731       1741       **SKQDDQQPLD LEGVKRKMDQ GNYTSVLEFS DDIVKIIQAA INSDGGQPEI****eeeeeeeeee eeebeeeeee eebbbbbebb eebbebbbbb beeeeeeeeb**  **f**      **f ff**  **f** **ff**     **s** **ss**   **s ff**     **s** **s** **f** **fff** **ff**  |
| 1751       1761       1771       1781       1791       **KKANSMVKSF FIRQMERVFP WFSVKKSRFW EPNKVSSNSG MLPNAVLPPS****eebeeebebb bbeebeebbe bbebeebebe eeeeeeeeee bbbebbbbeb****ffsfff** **f**   **s** **ffsff** **sf**   **sffsf** **f f**  **ff**  **fff** **ssfssssfs** |
| 1801       1811       1821       1831       1841       **LDHNYAQWQE REENSHTEQP PLMKKIIPAP KPKGPGEPDS PTPLHPPTPP****beeebbebee eeeeeeeeee ebbeebbeee eeeeeeeeee eeeeeeeeee****sfffssfsff ff**      **ff**   **ffssfff**  **f**    **f** **f f**          |
| 1851       1861       1871       1881       1891       **ILSTDRSRED SPELNPPPGI EDNRQCALCL TYGDDSANDA GRLLYIGQNE****eeeeeeeeee eeeeeeeeee eeeeebbbbb eeeeeeeeeb bebbbbbeee**        **ff ff**   **fff**    **ffffs** **sss** **ffff**  **f**   **sfsss** **sfff** |
| 1901       1911       1921       1931       1941       **WTHVNCALWS AEVFEDDDGS LKNVHMAVIR GKQLRCEFCQ KPGATVGCCL****beebbbbbbb bebbeeeeee beebbbbbbe beebebebbe eeeeebbbbb****sff** **ssssss sfssffffff sffss** **ss** **f sff** **fsf** **s**    **fffs** **ss**  |
| 1951       1961       1971       1981       1991       **TSCTSNYHFM CSRAKNCVFL DDKKVYCQRH RDLIKGEVVP ENGFEVFRRV****eeeeeebebb bbebeebbbb eeeeeebeeb eebbeeebee eebbebbeeb****ffffffsfss ssf**   **s** **ss ffffffsffs** **f**  **fff** **f**    **ssfs** **ff**  |
| 2001       2011       2021       2031       2041       **FVDFEGISLR RKFLNGLEPE NIHMMIGSMT IDCLGILNDL SDCEDKLFPI****bbbbebbbbe eebbeebeee ebbbbbbbbb bbbbbbbeeb eebeeebbeb** **ss** **fs**   **f ff** **s** **fsfff**  **s** **ssssss** **ssss** **s** **fs ffsf** **fssf**  |
| 2051       2061       2071       2081       2091       **GYQCSRVYWS TTDARKRCVY TCKIVECRPP VVEPDINSTV EHDENRTIAH****bbebbebebb beebeeebbb ebebbebeee eeeeebeeeb eeeeebbbee****ssfssfsfss sffsfffsss fsfs** **f**  **f**      **f** **fff**   **f**  **f** **s**  **f** |
| 2101       2111       2121       2131       2141       **SPTSFTESSS KESQNTAEII SPPSPDRPPH SQTSGSCYYH VISKVPRIRT****beeebeeeee eeeeeeeeee beeeeeeeee eeeeeebebe ebeeeeebeb****s**                                                 **f**    |
| 2151       2161       2171       2181       2191       **PSYSPTQRSP GCRPLPSAGS PTPTTHEIVT VGDPLLSSGL RSIGSRRHST****eeeeeeeeee eeeeeeeeee eeeeeeebbb bbbebbeebb ebbbbeeeee**       **ff**   **ffffffff**       **ff** **ss s** **sf** **sffs**       **ffff**  |
| 2201       2211       2221       2231       2241       **SSLSPQRSKL RIMSPMRTGN TYSRNNVSSV STTGTATDLE SSAKVVDHVL****eeeeeeeeeb ebbbeeeeee eebeeeeeeb eeeeeeeeee eeeeeeeebe****ff**                                                     |
| 2251       2261       2271       2281       2291       **GPLNSSTSLG QNTSTSSNLQ RTVVTVGNKN SHLDGSSSSE MKQSSASDLV****eeeeeeeeee eeeeeeeeee ebeeeeeeee eeeeeeeeee eeeeeeeebb**                                                       |
| 2301       2311       2321       2331       2341       **SKSSSLKGEK TKVLSSKSSE GSAHNVAYPG IPKLAPQVHN TTSRELNVSK****eeeeeeeeee eeeeeeeeee eebebbeeee eeeeeeeeee eeeeeeeeee**         **f**                                           28 |
| 2351       2361       2371       2381       2391       **IGSFAEPSSV SFSSKEALSF PHLHLRGQRN DRDQHTDSTQ SANSSPDEDT****eeebeeeeee eeeeeeeeeb eebeeeeeee eeeeeeeeee eeeeeeeeee**                                                       |
| 2401       2411       2421       2431       2441       **EVKTLKLSGM SNRSSIINEH MGSSSRDRRQ KGKKSCKETF KEKHSSKSFL****eeeebeeeee eeebeebeee beeeeeeeee eeeeeeeeeb eeeeeeeebe**                                                       |
| 2451       2461       2471       2481       2491       **EPGQVTTGEE GNLKPEFMDE VLTPEYMGQR PCNNVSSDKI GDKGLSMPGV****eeeeeeeeee eeeeeebeee bbeeeebeee bbeeeeeeee eeeeeeeeee**                                                       |
| 2501       2511       2521       2531       2541       **PKAPPMQVEG SAKELQAPRK RTVKVTLTPL KMENESQSKN ALKESSPASP****eeeeeeeeee eeeeeeeeee eebebebeeb eeeeeeeeee eeeeeeeeee**                    **f**       **ff**                        |
| 2551       2561       2571       2581       2591       **LQIESTSPTE PISASENPGD GPVAQPSPNN TSCQDSQSNN YQNLPVQDRN****eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee eeebeeeeee**                                                       |
| 2601       2611       2621       2631       2641       **LMLPDGPKPQ EDGSFKRRYP RRSARARSNM FFGLTPLYGV RSYGEEDIPF****bbbeeeeeee eeeeeeeebe eeebebebeb bbbbbbbbbb eeeeeeebeb**       **f**   **f**  **f** **fffsf fffsfsfsfs ssssss** **sss fffffff** **fs** |
| 2651       2661       2671       2681       2691       **YSSSTGKKRG KRSAEGQVDG ADDLSTSDED DLYYYNFTRT VISSGGEERL****eeeeeeeeee eeeeeeebee beeeeeeeee ebbbbbbbbb ebbeeeeeeb** **f** **f** **ffff**  **fffffffsff sff** **ffff**   **f** **sssss** **ss f**       **f**  |
| 2701       2711       2721       2731       2741       **ASHNLFREEE QCDLPKISQL DGVDDGTESD TSVTATTRKS SQIPKRNGKE****eeeebbeeee eeebeebeeb eeeeeeeeee eebebbeeeb eeeeeeeeee**     **sf** **f**        **sffs fffff** **ffff** **f**             **ff**  **ff** |
| 2751       2761       2771       2781       2791       **NGTENLKIDR PEDAGEKEHV TKSSVGHKNE PKMDNCHSVS RVKTQGQDSL****eeeeebeeee eeeeeeeeeb beeeeeeeee eeeeebeeeb ebeeeeeebb****ff**               **f**           **f**    **f**  **f** **f**      **f**   **ff**   |
| 2801       2811       2821       2831       2841       **EAQLSSLESS RRVHTSTPSD KNLLDTYNTE LLKSDSDNNN SDDCGNILPS****eeebeeeeeb eebeeeeeee eebbeebbee bbeeeeeeee eeebeebbbb**  **f** **fff**    **ff** **f**   **f**   **f**   **f**      **s** **ffffffff fff** **ffssss** |
| 2851       2861       2871       2881       2891       **DIMDFVLKNT PSMQALGESP ESSSSELLNL GEGLGLDSNR EKDMGLFEVF****ebbbbbbeee eebebeeeee eeeeeebbeb beebeeeeee eeebbbbebb****fssssssfff ffsf** **fffff ffffffs**  **s**  **f** **fff** **ff** **ff**  **ssfss** |
| 2901       2911       2921       2931       2941       **SQQLPTTEPV DSSVSSSISA EEQFELPLEL PSDLSVLTTR SPTVPSQNPS****eeebeeeeee eeebbbbbee eeebebeeeb eeebebbbee eeeeeeeeee****fffs**   **f** **f ff** **s**    **f**  **ff**  **fsfffs fffsfsssff ff** **f**  **ff**  29 |
| 2951       2961       2971       2981       2991       **RLAVISDSGE KRVTITEKSV ASSESDPALL SPGVDPTPEG HMTPDHFIQG****ebbbbeeeee eebebeeeee eeeeeeebbb eeeeeeeeee ebeeeebbeb**    **s**      **ff**     **f**                         **f** **fff** **s**    |
| 3001       3011       3021       3031       3041       **HMDADHISSP PCGSVEQGHG NNQDLTRNSS TPGLQVPVSP TVPIQNQKYV****bbebeebeee eeeeeeeeee eeeebeeeee eeebebeeee ebeeeeeeee**     **f**  **ff**             **f**   **f**    **ff**    **ffff**       **ff**  |
| 3051       3061       3071       3081       3091       **PNSTDSPGPS QISNAAVQTT PPHLKPATEK LIVVNQNMQP LYVLQTLPNG****eeeeeeeeee ebbeebbeeb eeebeebeee bbbbeeebee bebbeebeee**       **f** **f**       **f**     **f** **f**    **f**  **ssfff** **ff sfs** **ff** **fff** |
| 3101       3111       3121       3131       3141       **VTQKIQLTSS VSSTPSVMET NTSVLGPMGG GLTLTTGLNP SLPTSQSLFP****beeebebeee eeebeebbee ebebbbbbbe bbbbbeeeee eeeeeeeebe** **fffsf** **f**   **f**       **f**      **s**         **s**              **f**   |
| 3151       3161       3171       3181       3191       **SASKGLLPMS HHQHLHSFPA ATQSSFPPNI SNPPSGLLIG VQPPPDPQLL****eebeebbebe eeebbbebee beeeeeeeee eeeeebbbbb beeeeeeebb**   **f**       **f**      **s**        **f**           **s** **ss sf**     **f** **s** |
| 3201       3211       3221       3231       3241       **VSESSQRTDL STTVATPSSG LKKRPISRLQ TRKNKKLAPS STPSNIAPSD****beeeeeeeee eeeeeeeeee beeeebeebe eeeeeeeeee eeeeebbbee** **ff**   **f**                **fff** **sffs**   **f**  **ff** **fff**           |
| 3251       3261       3271       3281       3291       **VVSNMTLINF TPSQLPNHPS LLDLGSLNTS SHRTVPNIIK RSKSSIMYFE****bbbbbbbbbb beeebeeeee bbebbbbeee eeeebeebbe ebeeebbbbe**   **sss**  **s**     **f**   **f**     **fs**        **ff**  **ffssf f** **f**    **ssf** |
| 3301       3311       3321       3331       3341       **PAPLLPQSVG GTAATAAGTS TISQDTSHLT SGSVSGLASS SSVLNVVSMQ****ebeebeeeee eebbebeeee ebbeeeeebe ebebebbbee eebbbbbbbe**   **f** **f**                                      **f**  **sss**     |
| 3351       3361       3371       3381       3391       **TTTTPTSSAS VPGHVTLTNP RLLGTPDIGS ISNLLIKASQ QSLGIQDQPV****beeeeeeeee eeeebbbebe ebbbebeebb bbebbbebbe eeeebeeeee**                 **s**                  **ss** **f**    **f** **f**        |
| 3401       3411       3421       3431       3441       **ALPPSSGMFP QLGTSQTPST AAITAASSIC VLPSTQTTGI TAASPSGEAD****ebbeeeeebb eeeeeeeeee bbbbbbebbb bbeeeeebeb ebeeeeeeee**                                   **f**                   |
| 3451       3461       3471       3481       3491       **EHYQLQHVNQ LLASKTGIHS SQRDLDSASG PQVSNFTQTV DAPNSMGLEQ****eeeebeebee bbeeeeeeee eeeeeeeeee eeeeebeebb eeeeeeeeee**                                                       |
| 3501       3511       3521       3531       3541       **NKALSSAVQA SPTSPGGSPS SPSSGQRSAS PSVPGPTKPK PKTKRFQLPL****eeeeeeebeb eeeeeeebeb eeeeeeeeee eeeeeeeeee eebeebeeee**                      **f**                   **f** **f** **ff**      |
| 3551       3561       3571       3581       3591       **DKGNGKKHKV SHLRTSSSEA HIPDQETTSL TSGTGTPGAE AEQQDTASVE****eeeeeeeeee eebeeeeeee ebeeeeeeee eeeeeeeeee eeeeeeeeee**     **ffff**                                             30 |
| 3601       3611       3621       3631       3641       **QSSQKECGQP AGQVAVLPEV QVTQNPANEQ ESAEPKTVEE EESNFSSPLM****eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeeeeeeee eeeebebebb**                                                 **f** **f**   |
| 3651       3661       3671       3681       3691       **LWLQQEQKRK ESITEKKPKK GLVFEISSDD GFQICAESIE DAWKSLTDKV****bbbeeeeeee eeeeeeeeee ebbbebbeee bbebbbebbe ebbeebeeeb**     **f**              **f** **s** **sfs** **fff ssfs**  **f**  **f** **ssffsfffs** |
| 3701       3711       3721       3731       3741       **QEARSNARLK QLSFAGVNGL RMLGILHDAV VFLIEQLSGA KHCRNYKFRF****eebebeeebe ebbbebbbbb bbbbbbbebb bbbbeebeeb eebeebebeb****ffsfsfffsf f** **ssfss** **s**   **sss**  **sfss** **ss** **ffs** **fs** **fs** **fs** **sfs** |
| 3751       3761       3771       3781       3791       **HKPEEANEPP LNPHGSARAE VHLRKSAFDM FNFLASKHRQ PPEYNPNDEE****eeeeeeeeee eeeeeebebe bebeebbbeb bebbbeeeee eeeeeeeeee****f** **f** **f**    **f** **fffffsfsf** **f** **f** **s** **sfs sfsssfffff ff** **f** **f**  **ff** |
| 3801       3811       3821       3831       3841       **EEEVQLKSAR RATSMDLPMP MRFRHLKKTS KEAVGVYRSP IHGRGLFCKR****eeeeebeebe ebebeebebe bbbeebeeeb eebbebbebb bebebbbbee****fffffsffsf f** **fsffsf**    **ssffs**  **fs ff** **sfs** **fs**  **sfsfssssff** |
| 3851       3861       3871       3881       3891       **NIDAGEMVIE YAGNVIRSIQ TDKREKYYDS KGIGCYMFRI DDSEVVDATM****ebebbebbbe bbbebbbbbb eeeeeeeeee eebbbbbbbb eeeebbbbbb****fsfssfsssf sssfssss** **s fffffffff**  **ffsss** **ssss ff** **fsssss**  |
| 3901       3911       3921       3931       3941       **HGNAARFINH SCEPNCYSRV INIDGQKHIV IFAMRKIYRG EELTYDYKFP****bbbbbebbeb ebeeebbbeb bebeeeebbb bbbbbebeee eebebebebe****sssssfssfs f** **fffs** **sfs**   **ffffsss sss** **sfs** **ff ffsfsfsfsf** |
| 3951       3961       3971       **IEDASNKLPC NCGAKKCRKF LN****beebeeebeb eeeeeebeee ee****sffs** **ffs** **s ffffffsfff ff** |



**Figure S2**. Evolutionary conservation analysis performed for *MLL1* protein sequence 1 to 3969 aa using ConSurf.Amino acids are shaded according to their conservation ratings and levels. A grade of 1 represents fast-evolving (variable) sites and is marked in turquoise; a grade of 5 denotes sites evolving at a standard rate, shown in white; while a grade of 9 points to (evolutionarily conserved) sites that evolve slowly, represented in maroon. When a particular position covers a range of 4 or less color grades, its score is viewed as unreliable. These positions are highlighted in light yellow in the visual display.

**Table S1.** Predicted Functional Partners obtained from the STRING server.

54

****

|  |  |
| --- | --- |
| **Score** | **Predicted Functional Partners:** |
| 0.999 | Retinoblastoma-binding protein 5; In embryonic stem (ES) cells, plays a crucial role in the differentiation potential, particularly along the neural lineage, regulating gene induction and H3 'Lys-4' methylation at key developmental loci, including that mediated by retinoic acid (By similarity). As part of the MLL1/MLL complex, involved in mono-, di- and trimethylation at 'Lys-4' of histone H3. Histone H3 'Lys-4' methylation represents a specific tag for epigenetic transcriptional activation; WD repeat domain-containing | RBBP5 |
| 0.999 | Menin; Essential component of a MLL/SET1 histone methyltransferase (HMT) complex, a complex that specifically methylates 'Lys-4' of histone H3 (H3K4). Functions as a transcriptional regulator. Binds to the TERT promoter and represses telomerase expression. Plays a role in TGFB1-mediated inhibition of cell-proliferation, possibly regulating SMAD3 transcriptional activity. Represses JUND-mediated transcriptional activation on AP1 sites and that mediated by NFKB subunit RELA. Positively regulates HOXC8 and HOXC6 gene expression. May be involved in normal hematopoiesis through the [...] | MEN1 |
| 0.999 | Set1/Ash2 histone methyltransferase complex subunit ASH2; Component of the Set1/Ash2 histone methyltransferase (HMT) complex, a complex that specifically methylates 'Lys-4' of histone H3, but not if the neighboring 'Lys-9' residue is already methylated. As part of the MLL1/MLL complex it is involved in methylation and dimethylation at 'Lys-4' of histone H3. May function as a transcriptional regulator. May play a role in hematopoiesis; PHD finger proteins | ASH2L |
| 0.999 | WD repeat-containing protein 5; Contributes to histone modification. May position the N- terminus of histone H3 for efficient trimethylation at 'Lys-4'. As part of the MLL1/MLL complex it is involved in methylation and dimethylation at 'Lys-4' of histone H3. H3 'Lys-4' methylation represents a specific tag for epigenetic transcriptional activation. As part of the NSL complex it may be involved in acetylation of nucleosomal histone H4 on several lysine residues. May regulate osteoblasts differentiation; Belongs to the WD repeat WDR5/wds family | WDR5 |
| 0.992 | CREB-binding protein; Acetylates histones, giving a specific tag for transcriptional activation. Also acetylates non-histone proteins, like NCOA3 and FOXO1. Binds specifically to phosphorylated CREB and enhances its transcriptional activity toward cAMP-responsive genes. Acts as a coactivator of ALX1. Acts as a circadian transcriptional coactivator which enhances the activity of the circadian transcriptional activators: NPAS2-ARNTL/BMAL1 and CLOCK- ARNTL/BMAL1 heterodimers. Acetylates PCNA; acetylation promotes removal of chromatin-bound PCNA and its degradation during nucleotide excisi [...] | CREBBP |
| 0.988 | Host cell factor 1; Involved in control of the cell cycle. Also antagonizes transactivation by ZBTB17 and GABP2; represses ZBTB17 activation of the p15(INK4b) promoter and inhibits its ability to recruit p300. Coactivator for EGR2 and GABP2. Tethers the chromatin modifying Set1/Ash2 histone H3 'Lys-4' methyltransferase (H3K4me) and Sin3 histone deacetylase (HDAC) complexes (involved in the activation and repression of transcription, respectively) together. Component of a THAP1/THAP3-HCFC1-OGT complex that is required for the regulation of the transcriptional activity of RRM1. As part o [...] | HCFC1 |
| 0.980 | Protein dpy-30 homolog; As part of the MLL1/MLL complex, involved in the methylation of histone H3 at 'Lys-4', particularly trimethylation. Histone H3 'Lys-4' methylation represents a specific tag for epigenetic transcriptional activation. May play some role in histone H3 acetylation. In a teratocarcinoma cell, plays a crucial role in retinoic acid-induced differentiation along the neural lineage, regulating gene induction and H3 'Lys-4' methylation at key developmental loci. May also play an indirect or direct role in endosomal transport; Belongs to the dpy-30 family | DPY30 |
| 0.969 | PC4 and SFRS1-interacting protein; Transcriptional coactivator involved in neuroepithelial stem cell differentiation and neurogenesis. Involved in particular in lens epithelial cell gene regulation and stress responses. May play an important role in lens epithelial to fiber cell terminal differentiation. May play a protective role during stress-induced apoptosis. Isoform 2 is a more general and stronger transcriptional coactivator. Isoform 2 may also act as an adapter to coordinate pre-mRNA splicing. Cellular cofactor for lentiviral integration; Heparin binding growth factor family | PSIP1 |
| 0.969 | Chromodomain-helicase-DNA-binding protein 8; DNA helicase that acts as a chromatin remodeling factor and regulates transcription. Acts as a transcription repressor by remodeling chromatin structure and recruiting histone H1 to target genes. Suppresses p53/TP53-mediated apoptosis by recruiting histone H1 and preventing p53/TP53 transactivation activity. Acts as a negative regulator of Wnt signaling pathway by regulating beta-catenin (CTNNB1) activity. Negatively regulates CTNNB1- targeted gene expression by being recruited specifically to the promoter regions of several CTNNB1 responsiv [...] | CHD8 |
| 0.963 | Transcriptional activator Myb; Transcriptional activator; DNA-binding protein that specifically recognize the sequence 5'-YAAC[GT]G-3'. Plays an important role in the control of proliferation and differentiation of hematopoietic progenitor cells; Myb/SANT domain containing | MYB |

**Table S2.** GO enrichment analysis of the STRING networks created using the core module memberships revealed unique functions in the aging-associated module.

57

**Table S2.**Functional enrichments analysis of *MLL1* network



|  |
| --- |
|  **2.1.Biological Process (Gene Ontology)** |
| **False discovery rate** | **Strength** | **Count in network** | **Description** | **GO-term** |
| 5.47e-05 | 2.58 | 3 of [14](https://string-db.org/cgi/network?network_term_id=GO:0043984&input_query_species=9606) | Histone h4-k16 acetylation | [GO:0043984](http://amigo.geneontology.org/amigo/term/GO%3A0043984) |
| 0.00010 | 2.47 | 3 of [18](https://string-db.org/cgi/network?network_term_id=GO:0051571&input_query_species=9606) | Positive regulation of histone h3-k4 methylation | [GO:0051571](http://amigo.geneontology.org/amigo/term/GO%3A0051571) |
| 0.0063 | 2.47 | 2 of [12](https://string-db.org/cgi/network?network_term_id=GO:0019042&input_query_species=9606) | Viral latency | [GO:0019042](http://amigo.geneontology.org/amigo/term/GO%3A0019042) |
| 1.76e-06 | 2.39 | 4 of [29](https://string-db.org/cgi/network?network_term_id=GO:1904837&input_query_species=9606) | beta-catenin-TCF complex assembly | [GO:1904837](http://amigo.geneontology.org/amigo/term/GO%3A1904837) |
| 1.90e-08 | 2.38 | 5 of [37](https://string-db.org/cgi/network?network_term_id=GO:0051568&input_query_species=9606) | Histone h3-k4 methylation | [GO:0051568](http://amigo.geneontology.org/amigo/term/GO%3A0051568) |
| 0.0092 | 2.35 | 2 of [16](https://string-db.org/cgi/network?network_term_id=GO:0043982&input_query_species=9606) | Histone h4-k8 acetylation | [GO:0043982](http://amigo.geneontology.org/amigo/term/GO%3A0043982) |
| 0.0092 | 2.35 | 2 of [16](https://string-db.org/cgi/network?network_term_id=GO:0043981&input_query_species=9606) | Histone h4-k5 acetylation | [GO:0043981](http://amigo.geneontology.org/amigo/term/GO%3A0043981) |
| 2.96e-09 | 2.18 | 6 of [71](https://string-db.org/cgi/network?network_term_id=GO:0034968&input_query_species=9606) | Histone lysine methylation | [GO:0034968](http://amigo.geneontology.org/amigo/term/GO%3A0034968) |
| 3.06e-10 | 2.13 | 7 of [93](https://string-db.org/cgi/network?network_term_id=GO:0016571&input_query_species=9606) | Histone methylation | [GO:0016571](http://amigo.geneontology.org/amigo/term/GO%3A0016571) |
| 0.0233 | 2.09 | 2 of [29](https://string-db.org/cgi/network?network_term_id=GO:0030511&input_query_species=9606) | Positive regulation of transforming growth factor beta receptor signaling pathway | [GO:0030511](http://amigo.geneontology.org/amigo/term/GO%3A0030511) |
| 6.41e-07 | 2.04 | 5 of [82](https://string-db.org/cgi/network?network_term_id=GO:0045652&input_query_species=9606) | Regulation of megakaryocyte differentiation | [GO:0045652](http://amigo.geneontology.org/amigo/term/GO%3A0045652) |
| 0.00015 | 1.83 | 4 of [105](https://string-db.org/cgi/network?network_term_id=GO:0016573&input_query_species=9606) | Histone acetylation | [GO:0016573](http://amigo.geneontology.org/amigo/term/GO%3A0016573) |
| 1.36e-09 | 1.68 | 8 of [299](https://string-db.org/cgi/network?network_term_id=GO:0018205&input_query_species=9606) | Peptidyl-lysine modification | [GO:0018205](http://amigo.geneontology.org/amigo/term/GO%3A0018205) |
| 3.42e-09 | 1.61 | 8 of [351](https://string-db.org/cgi/network?network_term_id=GO:0016570&input_query_species=9606) | Histone modification | [GO:0016570](http://amigo.geneontology.org/amigo/term/GO%3A0016570) |
| 1.85e-06 | 1.61 | 6 of [260](https://string-db.org/cgi/network?network_term_id=GO:0045637&input_query_species=9606) | Regulation of myeloid cell differentiation | [GO:0045637](http://amigo.geneontology.org/amigo/term/GO%3A0045637) |
| 0.0298 | 1.41 | 3 of [206](https://string-db.org/cgi/network?network_term_id=GO:0006338&input_query_species=9606) | Chromatin remodeling | [GO:0006338](http://amigo.geneontology.org/amigo/term/GO%3A0006338) |
| 3.06e-10 | 1.4 | 10 of [713](https://string-db.org/cgi/network?network_term_id=GO:0006325&input_query_species=9606) | Chromatin organization | [GO:0006325](http://amigo.geneontology.org/amigo/term/GO%3A0006325) |
| 1.76e-06 | 1.4 | 7 of [493](https://string-db.org/cgi/network?network_term_id=GO:1903706&input_query_species=9606) | Regulation of hemopoiesis | [GO:1903706](http://amigo.geneontology.org/amigo/term/GO%3A1903706) |
| 0.00039 | 1.0 | 7 of 1253 | Positive regulation of transcription by rna polymerase ii | [GO:0045944](http://amigo.geneontology.org/amigo/term/GO%3A0045944) |
| 0.0147 | 1.0 | 5 of [895](https://string-db.org/cgi/network?network_term_id=GO:0000122&input_query_species=9606) | Negative regulation of transcription by rna polymerase ii | [GO:0000122](http://amigo.geneontology.org/amigo/term/GO%3A0000122) |
| 0.00014 | 0.95 | 8 of 1587 | Positive regulation of transcription, dna-templated | [GO:0045893](http://amigo.geneontology.org/amigo/term/GO%3A0045893) |
| 0.0069 | 0.92 | 6 of 1293 | Protein-containing complex assembly | [GO:0065003](http://amigo.geneontology.org/amigo/term/GO%3A0065003) |
| 3.89e-05 | 0.9 | 9 of 2005 | Positive regulation of cellular biosynthetic process | [GO:0031328](http://amigo.geneontology.org/amigo/term/GO%3A0031328) |
| 0.00033 | 0.88 | 8 of 1874 | Regulation of cell differentiation | [GO:0045595](http://amigo.geneontology.org/amigo/term/GO%3A0045595) |
| 0.0147 | 0.84 | 6 of 1528 | Negative regulation of nucleobase-containing compound metabolic process | [GO:0045934](http://amigo.geneontology.org/amigo/term/GO%3A0045934) |
| 0.00086 | 0.82 | 8 of 2172 | Regulation of transcription by rna polymerase ii | [GO:0006357](http://amigo.geneontology.org/amigo/term/GO%3A0006357) |
| 0.0171 | 0.71 | 7 of 2429 | Negative regulation of nitrogen compound metabolic process | [GO:0051172](http://amigo.geneontology.org/amigo/term/GO%3A0051172) |
| 0.0011 | 0.69 | 9 of 3239 | Positive regulation of nitrogen compound metabolic process | [GO:0051173](http://amigo.geneontology.org/amigo/term/GO%3A0051173) |
| 0.0267 | 0.68 | 7 of 2630 | Negative regulation of cellular metabolic process | [GO:0031324](http://amigo.geneontology.org/amigo/term/GO%3A0031324) |
| 0.0025 | 0.65 | 9 of 3600 | Positive regulation of macromolecule metabolic process | [GO:0010604](http://amigo.geneontology.org/amigo/term/GO%3A0010604) |
| 0.0116 | 0.64 | 8 of 3227 | Regulation of multicellular organismal process | [GO:0051239](http://amigo.geneontology.org/amigo/term/GO%3A0051239) |
| 0.0434 | 0.64 | 7 of 2875 | Negative regulation of macromolecule metabolic process | [GO:0010605](http://amigo.geneontology.org/amigo/term/GO%3A0010605) |
| 0.00030 | 0.56 | 11 of 5447 | Cellular component organization | [GO:0016043](http://amigo.geneontology.org/amigo/term/GO%3A0016043) |
| **2.2 Molecular Function (Gene Ontology)** |
|  |
| **False discovery rate** | **Strength** | **Count in network** | **Description** | **GO-term** |
| 0.0286 | 2.1 | 2 of [28](https://string-db.org/cgi/network?network_term_id=GO:0000217&input_query_species=9606) | DNA secondary structure binding | GO:0000217 |
| 0.0019 | 1.93 | 3 of [63](https://string-db.org/cgi/network?network_term_id=GO:0035064&input_query_species=9606) | Methylated histone binding | GO:0035064 |
| 0.0045 | 1.77 | 3 of [90](https://string-db.org/cgi/network?network_term_id=GO:0033613&input_query_species=9606) | Activating transcription factor binding | GO:0033613 |
| 0.00051 | 1.68 | 4 of [148](https://string-db.org/cgi/network?network_term_id=GO:0140030&input_query_species=9606) | Modification-dependent protein binding | GO:0140030 |
| 0.0019 | 1.51 | 4 of [221](https://string-db.org/cgi/network?network_term_id=GO:0042393&input_query_species=9606) | Histone binding | GO:0042393 |
| 0.00042 | 1.27 | 6 of [570](https://string-db.org/cgi/network?network_term_id=GO:0003682&input_query_species=9606) | Chromatin binding | GO:0003682 |
| 0.0445 | 1.1 | 4 of [571](https://string-db.org/cgi/network?network_term_id=GO:0003712&input_query_species=9606) | Transcription coregulator activity | GO:0003712 |
| 6.94e-05 | 1.09 | 8 of 1156 | Double-stranded dna binding | GO:0003690 |
| 0.00042 | 1.08 | 7 of 1028 | Transcription regulatory region sequence-specific dna binding | [GO:0000976](http://amigo.geneontology.org/amigo/term/GO%3A0000976) |
| 0.00042 | 0.81 | 9 of 2470 | DNA binding | GO:0003677 |

|  |
| --- |
| **2.3 Cellular Function (Gene Ontology)** |
| **False discovery rate** | **Strength** | **Count in network** | **Description** | **GO-term** |
| 4.04e-11 | 2.87 | 5 of [12](https://string-db.org/cgi/network?network_term_id=GO:0048188&input_query_species=9606) | Set1C/COMPASS complex | GO:0048188 |
| 3.10e-08 | 2.74 | 4 of [13](https://string-db.org/cgi/network?network_term_id=GO:0044666&input_query_species=9606) | MLL3/4 complex | GO:0044666 |
| 7.73e-12 | 2.54 | 6 of [31](https://string-db.org/cgi/network?network_term_id=GO:0071339&input_query_species=9606) | MLL1 complex | GO:0071339 |
| 1.85e-14 | 2.28 | 8 of [74](https://string-db.org/cgi/network?network_term_id=GO:0035097&input_query_species=9606) | Histone methyltransferase complex | GO:0035097 |
| 0.0016 | 1.81 | 3 of [83](https://string-db.org/cgi/network?network_term_id=GO:0000123&input_query_species=9606) | Histone acetyltransferase complex | GO:0000123 |
| 0.0049 | 1.6 | 3 of [133](https://string-db.org/cgi/network?network_term_id=GO:0034399&input_query_species=9606) | Nuclear periphery | GO:0034399 |
| 2.81e-09 | 1.33 | 9 of [752](https://string-db.org/cgi/network?network_term_id=GO:1990234&input_query_species=9606) | Transferase complex | [GO:1990234](http://amigo.geneontology.org/amigo/term/GO%3A1990234) |
| 4.63e-06 | 0.69 | 11 of 3973 | Nucleoplasm | GO:0005654 |

**2.4 KEGG Pathway**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Pathway** | **Description** | **Count in network** | **strength** | **False discovery rate** |
| hsa04934 | Cushing syndrome | 5 of 153 | 1.76 | 4.80e-0.6 |

|  |
| --- |
| **Table S3**. RegulomeDB variant classification scheme |
| **Description** | **Category** |
| Likely to affect binding and linked to the expression of a gene target |
| eQTL+ TF binding + matched TF motif + matched DNase footprint +DNase peak | 1a |
| eQTL+ TF binding + any motif + DNase footprint +DNase peak | 1b |
| eQTL+ TF binding + matched TF motif +DNase peak | 1c |
| eQTL+ TF binding + any motif + DNase footprint +DNase peak | 1d |
| eQTL+ TF binding + matched TF motif | 1e |
| eQTL+ TF binding / DNase peak | 1f |
| Likely to affect the binding |
| TF binding + matched TF motif + matched DNase footprint +DNase peak | 2a |
| TF binding + any motif +DNase footprint +DNase peak | 2b |
| TF binding + matched TF motif +DNase peak | 2c |
| Less likely to affect the binding |
| TF binding + any motif +DNase peak | 3a |
| TF binding + matched TF motif | 3b |
| Minimal binding evidence |
| TF binding +DNase peak | 4 |
| TF binding or DNase peak | 5 |
| Motif hit | 6 |