*Selection of potential candidates for new CLM biomarkers*

Among the genetic variants in 2010 SNPs detected in the genome wide sequencing analyses for the 10 patients, 75 genes for which the difference in the prevalence of the variants between the Group G and Group P had a P value of <0.2 were selected as preliminary candidates for new biomarkers (initial selection). Then, considering the incidence of mutations and the reported function of these 75 genes, the candidates for new biomarkers were narrowed down to the following 3 genes: *MICA, ARID 2* and *ZFHX3* (secondary selection) (**Supplemental Table 1**). Among these, *MICA* showed constant variants in a specific SNP (rs147557828), while the others showed indeterminate mutations or heterogeneous variants within the same gene (**Supplemental Table 2**). Therefore, *MICA*, the major histocompatibility complex (MHC) class I polypeptide-related sequence A, was selected for further study as a potential new biomarker for CLM.

|  |  |  |  |
| --- | --- | --- | --- |
| **Supplemental Table 1. Candidates for new biomarkers** | | | |
| Genes | Mutation in Group G (n=6) | Mutation in Group P (n=4) | Function |
| *NUMA1* | 1 | 4 | nuclear matrix |
| *NSD1* | 2 | 4 | androgen receptor transactivation |
| *BYSL* | 2 | 4 | rRNA processing |
| *NUMBL* | 0 | 3 | NFkB signaling |
| *MUC22* | 1 | 3 | unknown |
| *FOXO1* | 1 | 3 | transcription factor |
| ***MICA*** | **1** | **3** | **immune response** |
| *LOC114110* | 1 | 3 | unknown |
| ***ARID2*** | **1** | **3** | **chromatin organization** |
| *CCDC168* | 1 | 3 | unknown |
| *AGUSBP11* | 1 | 3 | unknown |
| ***ZFHX3*** | **1** | **3** | **tumor suppressor** |
| *A1BG* | 0 | 2 | unknown |
| *BRDT* | 0 | 2 | transcription regulation |
| *WDR60* | 0 | 2 | formation of cilia |
| *CENPP* | 0 | 2 | mitosis |
| *ATP10A* | 0 | 2 | cation transport ATPase |
| *LOC152217* | 0 | 2 | unknown |
| *CGNL1* | 0 | 2 | cell-cell junction |
| *ADAMTS2* | 0 | 2 | proteinase |
| *CLASRP* | 0 | 2 | unknown |
| *CA10* | 0 | 2 | hydration of CO2 |
| *OC90* | 0 | 2 | unknown |
| *CNDP2* | 0 | 2 | nonspecific dipeptidase |
| *ARID1A* | 0 | 2 | transcriptional activation |
| *CRIPAK* | 0 | 2 | negative regulator of PAK1 |
| *cELA3B* | 0 | 2 | elastase |
| *TBCC* | 0 | 2 | beta tubulin folding |
| *EIF2AK3* | 0 | 2 | mitochondorial function |
| *ERBB3* | 0 | 2 | EGFR family |
| *DNAH14* | 0 | 2 | motor protein |
| *HIVEP3* | 0 | 2 | unknown |
| *THBD* | 0 | 2 | activation of protein C |
| *KIAA0947* | 0 | 2 | unknown |
| *FOXD1* | 0 | 2 | unknown |
| *LIAA0284* | 0 | 2 | unknown |
| *IGFN1* | 0 | 2 | unknown |
| *ALDH3B1* | 0 | 2 | aldehyde dehydrogenase |
| *KIAA1377* | 0 | 2 | unknown |
| *HMCN1* | 0 | 2 | unknown |
| *PELO* | 0 | 2 | spermatogenesis |
| *MXRA5* | 0 | 2 | matrix remodeling |
| *CEP164* | 0 | 2 | assembly of primary cilia |
| *MYH7B* | 0 | 2 | heavy chain of myosin |
| *MREG* | 0 | 2 | unknown |
| *PRRC2B* | 0 | 2 | unknown |
| *SHISA9* | 0 | 2 | unknown |
| *FLJ43860* | 0 | 2 | unknown |
| *NES* | 0 | 2 | unknown |
| *NCAM1* | 0 | 2 | cell adhesion |
| *SLC13A1* | 0 | 2 | unknown |
| *SZT2* | 0 | 2 | superoxide dismutase |
| *GLB1* | 0 | 2 | galactosidase |
| *PLCD3* | 0 | 2 | phospholipase |
| *OPN4* | 0 | 2 | photoreceptive opsin protein |
| *TPTE* | 0 | 2 | signal transduction |
| *TRIM58* | 0 | 2 | unknown |
| *MUC16* | 0 | 2 | unknown |
| *TNN* | 0 | 2 | unknown |
| *CDH23* | 0 | 2 | cell adhesion |
| *TTBK1* | 0 | 2 | neuron-specific kinase |
| *TSC1* | 0 | 2 | stabilization of tuberin |
| *HLA-C* | 3 | 0 | HLA C |
| *FGFR1* | 3 | 0 | FGFR family |
| *KCNN3* | 3 | 0 | calcium channel |
| *PLAG1* | 3 | 0 | associated with pleomorphic adenoma |
| *BCL9* | 3 | 0 | associated with B cell lymphoma |
| *FAM104B* | 3 | 0 | unknown |
| *TCL1A* | 3 | 0 | associated with T cell leukemia |
| *ZNF568* | 3 | 0 | unknown |
| *DNAH1* | 3 | 0 | structural support of cilia |
| *PRDM1* | 4 | 0 | unknown |
| *ATIC* | 4 | 0 | purine biosynthesis |
| *KIAA1549* | 5 | 0 | unknown |