A. Relevance of Top-Ranking Genes for Leukemia

(1) KIAA0220
KIAA0220 codes for a PI3-kinase-related kinase SMG-1 like protein. Although the molecular function of the encoded protein is not yet known, its homology to PI3-kinase makes it an exciting pharmacological target. The dysregulation of the PI3-kinase signaling pathway has been implicated in multiple cancer types\(^{24}\), and pharmacological agents targeting this pathway are currently in clinical trials. Our ranking suggests that the protein encoded by KIAA0220 could possibly evolve as a similar target for the therapeutic management of leukemia.

(2) G-gamma globin
Higher levels of G-gamma globin have been reported in ALL\(^{46}\). Translation of both gamma and delta globin mRNAs is blocked by AZT, an anti-HIV drug which also inhibits the proliferation of leukemic cells\(^{48}\).

(3) Delta-globin
See (2) above.

(4) Brain-expressed HHCPA78 homolog
Although the brain-expressed HHCPA78 homolog has not been implicated in AML/ALL, it has been identified in leukemia cells. It is expected to be a homolog of thioredoxin interacting protein (Entrez Gene), which could possibly be involved in the conversion of post-mitotic cells to differentiating ones.

(5) Myeloperoxidase
Myeloperoxidase is an established marker for AML.\(^{a}\)

(6) Probable protein disulfide isomerase ER-60 precursor
A set of chaperone proteins that included protein disulfide isomerase were identified as interesting targets in a global profiling of the cell surface proteome of cancer cells\(^{41}\). Furthermore, in a clinical study of

\(^{a}\)Corresponding author.

\(^{a}\)Myeloperoxidase was actually included in our set of positive training genes. Its appearance in the ranked list is an artefact of the data set; in some instances, the data set has multiple occurrences of the same gene. Although we removed the specific occurrences of genes that we used in training and output a ranking only over the remaining genes, training genes with multiple occurrences can appear in the ranked list due to this artefact.
leukemia patients, levels of protein disulfide isomerase were shown to be distinctly altered, and correlated
with resistance to chemotherapy\textsuperscript{47}.

(7) NPM1 Nucleophosmin
Nucleophosmin (NPM), a nucleocytoplasmic shuttling protein with prominent nucleolar localization, reg-
ulates the ARF-p53 tumor-suppressor pathway. NPM is a characteristic feature of a large subgroup of
patients with AML\textsuperscript{9}.

(8) CD34
CD34 is over-expressed in AML and may be valuable in detecting minimal residual disease\textsuperscript{39}. In a meta-
analysis of 2483 patients, it was shown to be associated with a poor remission rate\textsuperscript{17}. It has also been
used to target drugs to leukemic cells\textsuperscript{4}.

(9) Elongation factor-1-beta
No link found.

(10) CD24
CD24 is expressed on a majority of B-lineage ALL\textsuperscript{21}, and on CD31+/CD33+ myeloid cells in the bone
marrow of children with AML\textsuperscript{9}. Furthermore, among ALL patients, a low CD24/CD45 antigen density
ratio has been associated with a good prognosis\textsuperscript{22}.

(11) 60S ribosomal protein L23
Ribosomal proteins have been implicated by multiple studies to be linked with different types of cancer\textsuperscript{2}. Al-
though the direct relationship between 60S ribosomal protein L23 and AML/ALL has not yet been
established, it was reported that pokeweed antiviral protein (PAP) and ricin A chains, which inactivate
60S subunits, could prevent the growth of leukemia cells in mice\textsuperscript{36}, indicating that this ribosomal protein
could emerge as a key therapeutic target in the management of leukemia.

(12) 5-aminolevulinic acid synthase
5-Aminolevulinic acid synthase (ALAS) is the first enzyme of the heme biosynthesis pathway\textsuperscript{43}. Al-
though ALAS has not been directly implicated in AML/ALL, heme enhances globin gene transcription
and is essential for globin translation (see (2) and (3) above). Furthermore, heme also seems to play a
role in regulating either synthesis or stability of hemoproteins, many of which have been implicated in
tumorigenesis\textsuperscript{35}.

(13) HLA class II histocompatibility antigen, DR alpha chain precursor
HLA-DR is a positive immunophenotyping marker in most AML cells. In a study carried out to investiga-
te the clinical significance of surface antigens in AML, the expression of HLA-DR was reported to be
associated with a lower remission rate\textsuperscript{12}.

(14) Epstein-Barr virus small RNA-associated protein
This gene encodes a cytoplasmic ribosomal protein that is a component of the 60S subunit (see (11)
above), and belongs to the L22E family of ribosomal proteins. Although this ribosomal protein has not
been implicated in the context of ALL/AML, one of the pseudogenes of this gene is fused to the acute
myeloid leukemia 1 (AML1) gene (Entrez Gene).

(15) HNRPA1 Heterogeneous nuclear ribonucleoprotein A1
The expression of heterogeneous nuclear ribonucleoparticle A1 and A2 proteins is elevated in a variety
of human cancers, and is lower or absent in normal tissues. Interestingly, the knock-down of the ribonu-
cleoprotein with RNA-interference was shown to induce apoptosis (cell death) in a variety of tumors,
suggesting that these could be developed as interesting therapeutic targets\textsuperscript{33}.

(16) Azurocidin
Azurocidin is known to be a marker for AML\textsuperscript{6}.

(17) Red cell anion exchanger (EPB3, AE1, Band 3)
A recent study has revealed that the red cell anion exchanger Band 3 is the therapeutic target of arsenic
trioxide, which has attracted attention as a treatment for acute promyelocytic leukemia\textsuperscript{10}. The exact role

\textsuperscript{b}Source: http://www.cancer.gov/cancerinfo/pdq/treatment/childAML/healthprofessional.
of this target remains to be resolved, but could emerge as a potential therapeutic target.

(18) **Topoisomerase II beta**
Topoisomerase inhibitors, including teniposide and etoposide, are currently being used to treat certain forms of leukemia.

(19) **HLA class I histocompatibility antigen, F alpha chain precursor**
No link found.

(20) **Probable G protein-coupled receptor LCR1 homolog**
This G protein-coupled receptor is related to LCR1/chemokine receptor-4 (CXCR-4), which is overexpressed in bone-marrow derived blasts, and is implicated in leukemic marrow infiltration. Although this homolog has not been studied in leukemia, the culture of AML cells with the CXCR-4 ligand, SDF-1, promoted their survival, whereas addition of neutralizing CXCR4 antibodies or SDF-1 antibodies significantly decreased it, suggesting that the probable G protein-coupled receptor LCR1 could be an interesting target for therapeutics.

(21) **HLA-SB alpha gene (class II antigen)**
No link found.

(22) **Int-6**
Int-6/eIF3-p48 has been identified as a human protein that binds to the human T-cell leukemia virus type I Tax oncoprotein. Although the role of Int-6/eIF3-p48 in human carcinogenesis is unknown at the present time, its expression is down-regulated in two of the most common forms of cancer in humans, namely breast and lung tumors.

(23) **Alpha-tubulin**
Tubulin, the protein component of microtubules (cytoskeletal elements that are important for mitotic spindle assembly and cell division), is a key molecular target for cancer therapy. Interestingly, alpha tubulin is phosphorylated in leukemic cells, in contrast to normal lymphocytes where it exists in a non-phosphorylated state, suggesting that it might play a role in progression of leukemia.

(24) **Terminal transferase**
Terminal transferase (TdT) is an established marker for ALL; it is expressed in over 95% of ALL cases.

(25) **Glycophorin B precursor**
The role of glycophorin in tumor malignancies is not yet well-understood. However, the over-expression of the proto-oncogene c-myc, which is implicated in leukemia and other cancers, has been shown to repress the expression of glycophorin.

**B. Relevance of Top-Ranking Genes for Colon Cancer**

(1) **26 kDa cell surface protein TAPA-1**
TAPA-1/DC1/CD81 has been implicated in the migration of endothelial cells, a key step in angiogenesis (growth of new blood vessels; increases tumor growth) and carcinoma (metastasis).

(2) **Id1**
Id helix-loop-helix (HLH) proteins function as regulators of cell growth and differentiation and, when over-expressed, can induce malignant transformation from normal to cancer cells. The expression of Id proteins in adenocarcinoma has been shown to be at least in part a consequence of loss of p53 function (p53 is a tumor-suppressing gene), and contributes to the uncontrolled proliferation of tumor cells. Id1 has also been implicated in tumor angiogenesis. Interestingly, Id genes are normally expressed at very low levels in adults, making them attractive new targets for anti-cancer drug design.

(3) **Cleavage and polyadenylation specificity factor**
No link found.

(4) **Interferon-inducible protein 9-27**
This 17 kDa membrane protein plays a key role in mediating the anti-proliferative effects of interferons, which have proven clinically effective as anti-tumor agents in a subset of cancer types. Furthermore,
silencing of this protein is also implicated in immortalisation\textsuperscript{19}, a key step towards tumorigenesis. Also see (18) below.

(5) **Nonspecific crossreacting antigen**

Nonspecific crossreacting antigen (NCA) is a major component of carcinoembryonic antigen (CEA), which is an important tumor marker. It has been shown using northern blot hybridization that NCA is expressed predominantly in cancerous tissues, making it a useful marker for colon cancer\textsuperscript{38}.

(6) **cAMP response element regulatory protein (CREB2)**

Also known as activating transcription factor 2 (ATF-2), CREB2 binds to cAMP response element (CRE) either as a homo-dimer, or as a hetero-dimer in conjunction with activator proteins (AP1), such as Jun, fos and ATF/CREB families, which regulate transcription in response to extracellular signals and have a decisive role in cell proliferation, tumorigenesis and apoptosis\textsuperscript{14}. ATF-2 mRNA is implicated in several types of human cancers, such as gastric, colon, pancreatic, and esophageal cancers\textsuperscript{44}. It has also been implicated in driving the expression of vascular endothelial growth factor (VEGF) under endoplasmic reticulum stress, which could further promote tumor angiogenesis.

(7) **MHC class I HLA-Bw58**

No link found.

(8) **Translational initiation factor 2 gamma subunit**

No link found.

(9) **Splicing factor (CC1.4)**

The protein encoded by this gene is an RNA binding protein, which is found in the nucleus and co-localizes with the core spliceosomal protein. Studies of a murine protein with high sequence similarity to this protein suggest that this protein may act as a transcriptional co-activator for JUN/AP1, which have been shown to play a dominant role in the oncogenic ras-induced transformation of human carcinoma cells\textsuperscript{50}.

(10) **Nucleolar protein (B23)**

Nucleophosmin (B23) is involved in ribosome biogenesis, and interacts with tumor suppressor proteins p53 and Rb\textsuperscript{20}. Levels of nucleophosmin have been reported to be up-regulated in many tumor types\textsuperscript{31}.

(11) **Lactate dehydrogenase-A (LDH-A)**

Lactate dehydrogenase (LDH) levels have been correlated with poor prognosis and with resistance to chemotherapy and radiotherapy in various cancers. LDH is over-expressed in colorectal cancer\textsuperscript{32}, and has also been implicated in mediating c-myc-induced transformation from normal to cancer cells\textsuperscript{40}.

(12) **Guanine nucleotide-binding protein G(OLF), alpha subunit**

Persistent activation of the G-protein (olf) has been shown to exert convergent signals through the rho kinase pathway to promote cellular invasion and survival in solid tumors during towards metastasis\textsuperscript{37}.

(13) **LI-cadherin**

Over-expression of LI-cadherin has been implicated in lymph node metastasis of gastro-intestinal cancer, which is closely related to colon cancer\textsuperscript{18}.

(14) **Lysozyme**

Colonic epithelium can produce lysozyme, and its expression is up-regulated in the dysplastic epithelium in adenomas and in invasive cancer cells\textsuperscript{51}.

(15) **Prolyl 4-hydroxylase beta-subunit and disulfide isomerase (P4HB)**

This protein possesses two different enzymatic functions depending on whether it is present in cells in monomer form (disulfide isomerase) or in the prolyl 4-hydroxylase tetramer form\textsuperscript{34}. Interestingly, the expression of prolyl-hydroxylase was shown to suppress hypoxia inducible factor-1-alpha activation and inhibit angiogenesis and growth of colon carcinoma\textsuperscript{8}.

(16) **Eukaryotic initiation factor 4AII**

Eukaryotic translation initiation factor, eIF4A, exists as a complex with cyclin-dependent kinases (CDKs). The CDK-eIF4A complex is abundant in actively proliferating and growing cells, but is absent from cells that have ceased dividing, indicating that this interaction could underlie the molecular
mechanism linking cell proliferation with translational control, which is altered in cancer progression. Interestingly, the CDK-eIF4A complex contains kinase activity that is sensitive to the CDK-specific inhibitor roscovitine, suggesting that this may be a lead compound for the treatment of colon cancer\(^\text{15}\).

(17) **HLA class I histocompatibility antigen**  
No link found.

(18) **Interferon-inducible protein 1-8D**  
Interestingly, both interferon-inducible protein 1-8D and 9-27 (see (4) above) have been postulated to mediate the link between interferon- and radiation-induced cell death\(^\text{5}\). Levels of the former are also up-regulated in tumor cells following the suppression of bcr-abl synthesis by siRNAs or tyrosine kinase activity by Glivec, a novel anti-cancer drug\(^\text{52}\). These studies indicate that this protein could be an interesting therapeutic target for inducing tumor cell death.

(19) **Very long chain acyl-CoA dehydrogenase**  
No link found.

(20) **Dipeptidase**  
Dipeptidase 1 has been used as a marker for colon cancer\(^\text{28}\).

(21) **Heat shock 27 kDa protein**  
Low molecular weight stress proteins such as heat shock protein 27 (hsp27) have been implicated in cellular processes potentially related to malignant transformation from normal to cancer cells\(^\text{29}\). Furthermore, increased expression of hsp27 has been shown to enhance the tumorigenicity of immunogenic colon carcinoma\(^\text{11}\).

(22) **Tyrosine-protein kinase receptor TIE-1 precursor**  
Protein tyrosine kinases (PTKs) are a major class of proto-oncogenes that are involved in tumor progression and angiogenesis. Positive immunohistochemical staining for tie-1 was observed in gastric adenocarcinoma tissues\(^\text{23}\). Furthermore, clinico-pathological studies have indicated that tie-1 kinase expression is inversely correlated with patients’ survival, indicating that tie-1 inhibitors could have major implications in colon cancer.

(23) **Mitochondrial import receptor MOM38**  
No link found.

(24) **Mitochondrial matrix protein P1 precursor**  
Also known as heat shock protein 60 (hsp60), it belongs to a group of proteins that typically modulate the cellular response to stress but are also implicated in the cell cycle, cell proliferation and differentiation. Altered expression of HSP has been reported for nearly all classes of tumors, and hsp60 specifically has been shown to be over-expressed in colon cancer\(^\text{16}\).

(25) **Eukaryotic initiation factor EIF-4A homolog**  
See (16) above.

References


