Supplementary table 2. The list of 33 recurrent fusions. Relevant gene roles as oncogene or tumor suppressor (TSG) were obtained from Cancer Gene Census. Potential role of the fusion partners in cancer was further investigated by literature searches.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Fusion** | **Number of patients (out 270)** | **5’ gene role** | **3’ gene role** | **literature evidence** |
| LMO7:EXT2 | 30 |  | TSG | LMO7 and LIMCH1 interact with LRIG proteins in lung cancer, with prognostic implications for early-stage disease (PMID: 30429017) |
| CD27-AS1:MANBAL | 12 |  |  | Unknown |
| GABPB1:LMCD1 | 10 |  |  | Identification of a novel lncRNA GABPB1-IT1 that is downregulated and predicts a poor prognosis in non-small cell lung cancer (PMID: 31289561) |
| TTLL12:RAB17 | 10 |  |  | Next, utilizing RT-PCR and Sanger sequencing, we further confirmed the existence of such a new transcript isoform of TTLL12 in more human cancer cells including lung cancer cells and other cancer cells. (PMID: 27748896) |
| CSRNP2:PCSK7 | 6 |  |  | Patients with clear cell renal cell carcinoma where *CSRNP1* and *CSRNP3* had a low expression had worse overall survival (PMID: 33869003). PCSK7 activates HIF1A, a major transcription factor that promotes tumorigenesis under hypoxic conditions (PMID: 24436242). |
| ZNF587B:ZNF211 | 6 |  |  | The Novel Zinc Finger Protein 587B Gene, ZNF587B, Regulates Cell Proliferation and Metastasis in Ovarian Cancer Cells in vivo and in vitro (PMID: 32617020). |
| DNPEP:C9orf3 | 5 |  |  | DNPEP overexpression suppresses breast cancer cell proliferation and invasion and restricts breast cancer growth and metastasis in mice (PMID: 31219614). C9orf3 is a host gene of microRNA-23b/27b/24 cluster that facilitates colon cancer cell migration by targeting FOXP2 (PMID: 31936744). |
| ZNF267:ZNF720 | 5 |  |  | Zinc finger protein 267 is up-regulated in hepatocellular carcinoma and promotes tumor cell proliferation and migration (PMID: 21840307). |
| DPP9:BZRAP1 | 4 |  |  | Contribution of upregulated dipeptidyl peptidase 9 (DPP9) in promoting tumoregenicity, metastasis and the prediction of poor prognosis in non-small cell lung cancer (NSCLC) (PMID: 27943262). |
| PDE6D:CSGALNACT1 | 4 |  |  | The Delta Subunit of Rod-Specific Photoreceptor cGMP Phosphodiesterase (PDE6D) Contributes to Hepatocellular Carcinoma Progression (PMID: 30901922). |
| PDSS2:TRPV2 | 4 |  |  | Decrease of PDSS2 expression, a novel tumor suppressor, in non-small cell lung cancer (PMID: 23312889). |
| STX16:SIX4 | 4 |  |  | A homeobox family gene SIX4 was linked to lymph node metastasis and poor survival in NSCLC (PMID: 27821176). |
| CCDC57:AMOTL2 | 3 |  |  | Unknown |
| HSPG2:PLXNB1 | 3 |  |  | HSPG2 overexpression independently predicts poor survival in patients with acute myeloid leukemia (PMID: 32606327). |
| KIF13A:LINC00581 | 3 |  |  | Unknown |
| TLE3:AES | 3 |  |  | Targeting of TLE3 by miR-3677 in human breast cancer promotes cell proliferation, migration and invasion (PMID: 32002031). |
| ZNF669:ZNF136 | 3 |  |  | Unknown |
| ZNF331:ZNF529 | 2 | TSG |  | Zink finger protein 331 suppresses growth and invasiveness in gastric cancer (PMID: 22370639). |
| ATP13A4:MAN2B1 | 2 |  |  | ATP13A4-ALK fusions was detected in a case study of a single 53 year old non-smoker patient with adenocarcinoma (PMID: 32903930). |
| ATP2A2:IFT81 | 2 |  |  | Study suggests that germline alterations of ATP2A2 may predispose to lung and colon cancer and that an impaired ATP2A2 gene might be involved, directly or indirectly, as an early event in carcinogenesis (PMID: 17116488). |
| ATP5G2:SMARCC2 | 2 |  |  | Overexpression of LncRNA SMARCC2 enhances the proliferation and migration of gastric cancer cells, while inhibition of LncRNA SMARCC2 does the opposite (PMID: 29337109). |
| BAI1:BAI2 | 2 |  |  | BAI1 acts as a tumor suppressor in lung cancer A549 cells by inducing metabolic reprogramming via the SCD1/HMGCR module (PMID: 32255478). |
| MCFD2:DYNC1H1 | 2 |  |  | MCFD2 promotes cancer metastasis by regulating LMAN1 and LGALS3BP expression levels in oral squamous cell carcinoma (PMID: 29679592).  Missense mutations in DYNC1H1 are associated with various cancers (PMID: 28455235). |
| NOC4L:FBRSL1 | 2 |  |  | Unknown |
| NOXA1:SLCO4A1 | 2 |  |  | LncRNA SLCO4A1-AS1 drove NSCLC progression by activating NF-κB signaling pathway via sponging miR-223-3p to enhance IKKα expression (PMID: 32687454). |
| OPLAH:ALKBH6 | 2 |  |  | Transcriptomics analysis suggested the four enzymes related to glutathione metabolism-CD13, GPX4, RRM2B, and OPLAH-as potential targets of cisplatin resistance in nonsmall cell lung cancer (PMID: 31373204). |
| RTCB:GAREML | 2 |  |  | Genistein retarded tumor growth in head and neck cancers through the elevation of miR-34a and suppression of RTCB (PMID: 32610494). |
| SLC38A7:AC144831.1 | 2 |  |  | SNAT7 (SLC39A7) is the primary permeation pathway for glutamine across the lysosomal membrane and it is required for growth of cancer cells in a low free-glutamine environment (PMID: 28416685). |
| SNRPD3:LYST | 2 |  |  | Silencing of SNRPD3 leads to cytotoxic alternative splicing switch in the proteasome subunit beta 3 mRNA in NSCLC (PMID: 32545483). |
| UBE2D2:MORF4L1 | 2 |  |  | Upregulated Circular RNA circ-UBE2D2 Predicts Poor Prognosis and Promotes Breast Cancer Progression by Sponging miR-1236 and miR-1287 (PMID: 31336316). MORF4L1 suppresses cell proliferation, migration and invasion by increasing p21 and E-cadherin expression in nasopharyngeal carcinoma (PMID: 30655767). |
| ZEB2:SEC13 | 2 |  |  | LINC01296/miR-141-3p/ZEB1-ZEB2 axis promotes tumor metastasis in NSCLC by enhancing epithelial-mesenchymal transition process (PMID: 33854632). |
| ZNF528:ZNF766 | 2 |  |  | Unknown |
| ZNF746:LINC01395 | 2 |  |  | Inhibition of ZNF746 suppresses invasion and epithelial to mesenchymal transition in H460 non-small cell lung cancer cells (PMID: 24145959). |

Supplementary table 3. List of genes with the highest number of separate fusion partners as a measure of promiscuity. Only the genes with at least 3 different fusion partners were included in the table. Recurrence refers to the number of samples where the gene was observed forming transcript fusions. Role in cancer columns indicates whether the fusion gene is a known oncogene or tumor suppressor gene (TSG) included in COSMIC Cancer Gene Census (CGC).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gene** | **Fusions** | **Promiscuity** | **Recurrence** | **Role in cancer** | **Literature evidence** |
| RNF213 | NUTM2A-AS1:RNF213; MAFG:RNF213; USP16:RNF213; HELZ2:RNF213 | 4 | 4 |  | RNF213 mutations in circulating DNA differentiate between early stage lung cancer from pulmonary nodules (PMID: 33200540). |
| ZBTB7A | DNAJC11:ZBTB7A; EIF4G2:ZBTB7A; NDST1:ZBTB7A; EXD3:ZBTB7A | 4 | 4 |  | ZBTB7A acts as tumor suppressor by regulating glycolysis. Reduced expression of ZBTB7A correlates with poor survival in colon cancer patients. ZBTB7A tumors show heightened sensitivity to glycolysis inhibition (PMID: 25184678). |
| CCDC91 | CCDC91:USP15; FTX:CCDC91; HIPK1:CCDC91 | 3 | 3 |  | Uknown |
| CERS6 | CERS6:SAP130; CERS6:MTX2; CERS6:IMPA2 | 3 | 3 |  | CERS6 is required for cell migration and metastasis in NSCLC (PMID: 32902157). |
| CHMP1A | TDRP:CHMP1A; MSR1:CHMP1A; CYP24A1:CHMP1A | 3 | 3 |  | CHMP1A acts as a tumor suppressor gene in that inhibits proliferation of renal cell carcinoma (PMID: 22261332). |
| CHST11 | CHST11:SUCLG2; CHST11:BICD1; CHST11:MAPKAPK5 | 3 | 3 |  | CHST11 plays a role in the induction of epithelial-to-mesenchymal transition and stem cell-like properties in breast cancer (PMID: 32344408). |
| HIF1A | CHIC2:HIF1A; HIF1A:PRKCH; HIF1A:TNPO2 | 3 | 3 |  | HIF1 is key factor involved in the metabolic switch in cancer tissues from oxidative phosphorylation to aerobic glycolysis (PMID: 33791987). |
| INPP4A | INPP4A:VWA3B; INPP4A:TTC3; INPP4A:RAPGEF1 | 3 | 3 |  | Lower nuclear localization of INPP4A correlates with increased cancerous growth, moreover constitutive expression of this gene leads to massive apoptosis (PMID: 30071275). |
| ITCH | RALY:ITCH; BCL2L1:ITCH; ITCH:ZNF217 | 3 | 3 |  | Circular RNA cir-ITCH promotes cell invasion and migration in osteosarcoma (PMID: 31960764). |
| MAD1L1 | MAD1L1:NBPF19; DNAAF5:MAD1L1; EIF3B:MAD1L1 | 3 | 3 |  | Gain of a region on 7p22.3, containing MAD1L1, is the most frequent event in small-cell lung cancer cell lines (PMID: 16130125). |
| MKL1 | MKL1:PVT1; MKL1:RHOT1; MKL1:MTFR1 | 3 | 3 |  | MKL1 potentiates lung cancer cell migration and invasion by epigenetically activating MMP9 transcription (PMID: 25746000). |
| MROH1 | MROH1:ZNF395; ABT1:MROH1; COL23A1:MROH1 | 3 | 3 |  | Unknown |
| N4BP2L2 | USP12:N4BP2L2; TRAPPC9:N4BP2L2; N4BP2L2:SH3BGRL | 3 | 3 |  | Unknown |
| NFIB | ENPP2:NFIB; CD302:NFIB; GPATCH2L:NFIB | 3 | 3 |  | NFIB promotes the migration and progression of kidney renal clear cell carcinoma by regulating PINK1 transcription (PMID: 33981484). |
| NPEPPS | AGPAT6:NPEPPS; ITFG3:NPEPPS; NPEPPS:SLC7A11 | 3 | 3 |  | Unknown |
| PCNX | ZFYVE28:PCNX; PCNX:SPTLC2; PCNX:SYNJ2BP-COX16 | 3 | 3 |  | PCNX-3' UTR promotes cell growth, proliferation and cell cycle progression, and suppresses apoptosis of lung cancer cells (PMID: 28789966). |
| PPP1R12C | PPP1R12C:SLC6A3; PPP1R12C:TNK2; PPP1R12C:SPPL2B | 3 | 3 |  | Unknown |
| PTPRF | PTPRF:CSMD2; ACSF3:PTPRF; ATXN2L:PTPRF | 3 | 3 |  | PTPRF as a novel tumor suppressor through deactivation of ERK1/2 signaling in gastric adenocarcinoma (PMID: 30464527). |
| RHOT1 | RHOT1:UTP6; SLFN5:RHOT1; MKL1:RHOT1 | 3 | 3 |  | Propofol suppresses non-small cell lung cancer tumorigenesis by regulation of circ-RHOT1/miR-326/FOXM1 axis (PMID: 33515563). |
| SVIL | LCOR:SVIL; SVIL:ZNF292; SVIL:MTPAP | 3 | 3 |  | SVIL promotes tumor angiogenesis in liver cancer (PMID: 32468064). |
| TBL1XR1 | TBL1XR1:COL14A1; NSUN3:TBL1XR1; TBL1XR1:KDM6A | 3 | 3 | Oncogene, TSG | TBL1XR1 is involved in c-Met-mediated tumorigenesis of human nonsmall cell lung cancer (PMID: 31243347). |
| TMCC1 | TMCC1:HMCES; TMCC1:CNBP; MIR205HG:TMCC1 | 3 | 3 |  | Unknown |
| TRAPPC9 | RSF1:TRAPPC9; TRAPPC9:N4BP2L2; CLK3:TRAPPC9 | 3 | 3 |  | Elevated NIBP/TRAPPC9 mediates tumorigenesis of cancer cells through NFκB signaling (PMID: 25704885). |
| ZBTB20 | CD96:ZBTB20; ZBTB20:BRAF; ATG3:ZBTB20 | 3 | 3 | Oncogene | Zinc finger protein ZBTB20 promotes cell proliferation in non-small cell lung cancer through repression of FoxO1 (PMID: 25311537). |

Supplementary table 4. Significant associations between clinicopathological covariates and recurrent fusions. The associations were detected using logistic regression taking into account age, gender, smoking status, histology and tumor recurrence.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Fusion** | **Covariate** | **Estimate** | **Std.Error** | **Z value** | **Pr(>|Z|)** | **Fusion counts** |
| CD27-AS1:MANBAL | Gender | -1.83 | 0.85 | -2.13 | 0.03 | Female: 9/134  Male: 2/113 |
| DNPEP:C9orf3 | Recurrence | 4.18 | 1.98 | 2.1 | 0.04 | Recurred: 3/100  No recurrence: 1/154 |
| GABPB1:LMCD1 | Histology | 3.12 | 1.41 | 2.2 | 0.03 | Large cell neuroendocrine: 1/4  Rest: 8/245 |
| LMO7:EXT2 | Histology | 1.57 | 0.72 | 2.16 | 0.03 | Adenosquamous carcinoma: 4/9  Rest: 22/223 |

Supplementary table 5. Modeling survival based on the occurrence of transcript fusions and individual fusion partners. The survival analysis was conducted using Cox proportional hazards regression model adjusted for age, gender, smoking status, histology and tumor recurrence.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Covariate | coef | exp(coef) | se(coef) | z | P-value/log-rank test |
| smoking(Never) | -0.81 | 0.44 | 0.32 | -2.54 | 0.01 |
| Histology(Large cell carcinoma) | 1.06 | 2.88 | 0.47 | 2.23 | 0.03 |
| **stageII** | **0.69** | **2** | **0.2** | **3.45** | **<0.001** |
| **recurrence** | **1.86** | **6.45** | **0.2** | **9.11** | **<0.001** |
| **MROH1** | **2.64** | **14.04** | **0.75** | **3.53** | **0 /<0.001 \*** |
| N4BP2L2 | 1.44 | 4.2 | 0.61 | 2.35 | 0.02/0.15 |
| TRAPPC9 | 1.43 | 4.18 | 0.61 | 2.35 | 0.02/0.23 |