

Functional Genomics Screen with Pooled shRNA Library and Gene Expression Profiling with Extracts of *Azadirachta indica* Identify Potential Pathways for Therapeutic Targets in Head and Neck Squamous Cell Carcinoma

Neeraja M Krishnan¹⁺, Hiroto Katoh²⁺, Vinayak Palve¹, Manisha Pareek¹, Reiko Sato², Shumpei Ishikawa^{2*} and Binay Panda^{1*}

¹Ganit Labs, Bio-IT Centre, Institute of Bioinformatics and Applied Biotechnology, Bangalore 560100, India

²Department of Genomic Pathology, Medical Research Institute, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo, 113-8510, Japan

⁺ equal contribution

Correspondence:

Dr. Binay Panda

binay@ganitlabs.in

and

Professor Shumpei Ishikawa

sish.gpat@mri.tmd.ac.jp

Abstract

Tumor suppression by the extracts of *Azadirachta indica* (neem) works via anti-proliferation, cell cycle arrest, and apoptosis, demonstrated previously using cancer cell lines and live animal models. However, very little is known about the molecular targets and pathways that the neem extracts and the associated compounds act through. Here, we address this using a genome-wide functional pooled shRNA screen on head and neck squamous cell carcinoma cell line treated with crude neem leaf extracts, known for their anti-tumorigenic activity. By analyzing differences in global clonal sizes of the shRNA-infected cells cultured under no treatment and treatment with neem leaf extract conditions, assayed using next-generation sequencing, we found 225 genes affected the cancer cell growth in the shRNA-infected cells treated with neem extract. Pathway enrichment analyses of whole-genome gene expression data from cells temporally treated with neem extract revealed important roles played by the TGF- β pathway and *HSF-1*-related gene network. Our results indicate that neem extract simultaneously affects various important molecular signaling pathways in head and neck cancer cells, some of which may be therapeutic targets for this devastating tumor.

Keywords: Functional genomics screen, pooled shRNA screen, gene expression profile analysis, drug target discovery, Neem (*Azadirachta indica* A. Juss), head and neck squamous cell carcinoma, TGF- β , *HSF-1*

41 **Introduction**

42 Despite previous knowledge on the use of the ethnobotanical database, evidence-based
 43 know-how and the mechanism(s) of action of many plant-derived compounds are currently lacking.
 44 *Azadirachta indica*, neem, is one such species that is unique, versatile and important. The neem tree
 45 is also one of the most intensively studied sources of natural products with plant-derived extracts
 46 showing anti-bacterial, anti-fungal, anti-viral, anti-inflammatory, anti-hyperglycaemic,
 47 immunomodulatory, anti-malarial, and anti-carcinogenic properties (reviewed in (Subapriya and
 48 Nagini, 2005). Past efforts on research on neem have mostly been focused on pre-clinical studies,
 49 concentrating mainly on two purified neem-derived compounds, nimbolide and azadirachtin, due to
 50 their pharmaceutically- and agriculturally-important properties, respectively. The emergence of
 51 genome-wide tools and techniques has opened up new avenues to pursue research through
 52 systematic studies on cancer drug target discovery using neem-derived metabolites. One such
 53 method is RNA interference. It is a powerful and proven system to perturb gene function in higher
 54 organisms by phenotypic screening, with potential ability to identify pathways and networks of
 55 genes involved in cancer (Tang et al., 2008; Blakely et al., 2011; Poell et al., 2011; Iorns et al., 2012;
 56 Mendes-Pereira et al., 2012; Poell et al., 2012; Singleton et al., 2013). Use of pooled genome-wide
 57 libraries of short hairpin RNA (shRNA) coupled with next-generation sequencing offers higher
 58 sensitivity and broader dynamic range to screen for biological activity and therefore has the
 59 potential to identify the pathways for drug actions through network-based approaches (Erler and
 60 Linding, 2010; 2012).

61 Head and neck squamous cell carcinoma (HNSCC) is the sixth leading cause of cancer
 62 worldwide (Jemal et al., 2011) with a 5-year survival of less than 50% (Mishra and Meherotra,
 63 2014). In the past, the ability of neem-derived compounds to stop cell proliferation and their
 64 mechanism(s) of action have been reported, including in HNSCC cell lines (Roy et al., 2007; Harish
 65 Kumar et al., 2009; Harish Kumar et al., 2010; Priyadarsini et al., 2010; Gupta et al., 2013; Hsieh et

al., 2015; Kasala et al., 2015; Patel et al., 2016; Chien et al., 2017; Kowshik et al., 2017). However, efforts towards finding novel drug targets and the pathways using genome-wide tools for neem-derived compounds is currently lacking. The current study combined genome-wide shRNA based functional screening and whole-genome gene expression assay using a head and neck cancer cell line, paving ways to characterize and understand the critical molecular pathways for the therapeutic properties of the neem leaf extract against HNSCC.

Materials and Methods

Neem plant extracts

We tested leaf, bark, fruit, dry seed and twigs from neem plants to make ethanolic extracts. First, the plant parts were collected fresh and thoroughly washed with MilliQ water followed by 70% ethanol and the process was repeated thrice with a final wash with MilliQ water before they were air dried under controlled condition for several days. Once the plant organs were dry, 250gm of each organ was measured and ground separately using a grinder, mixed with 500ml of ethanol and transferred to a glass bottle. The mixture in the glass bottle was rocked overnight to mix thoroughly and then centrifuged at 5000g for 10min. The supernatant was separated, and the residue at the bottom was again treated with 500ml of ethanol with overnight rocking at room temperature followed by centrifugation. The process was repeated three times, and the supernatant was pooled from all the three spins before proceeding to the next step. The ethanolic extract (referred as neem extract) was evaporated using an oven at 55°C till the materials became very viscous and gooey like and then kept at 4°C until further use.

Reconstitution of neem extracts

Neem extracts were dissolved in DMSO to make a stock solution, which was serially diluted further with Dulbecco's Modified Eagles' Media (DMEM) to make a final concentration of 2mg/ml

as working stock solution. The working stock solution was passed through a 0.22micron sterile filter and stored until further use. The filtered and sterile stock solution was further diluted with DMEM to make the final solution and was used for treatment.

Cell culture and IC₅₀ calculation

Multiple human HNSCC cell lines (UPCI:SCC029B and UPCI:SCC040, both gifts from Dr.Susan Gollin, University of Pittsburgh, PA, USA (White et al., 2007); UM-SCC47, a gift from Dr. Thomas Carey, University of Michigan, MI, USA (Brenner et al., 2010); HSC-3 and HSC-4, purchased from RIKEN, Japan (Momose et al., 1989) were tested with all neem extracts before HSC-4 was chosen for further study as the cells showed excellent growth inhibitory pattern in the presence of neem extracts. All the cells were maintained in DMEM supplemented with 10% FBS, 1X MEM non-essential amino acids solution and 1X penicillin/streptomycin mixture (Gibco) at 37⁰C with 5% CO₂ incubation.

For IC₅₀ calculation, the xCELLigence Real-Time Cell Analysis (RTCA) DP instrument (Acea Bio, USA) was used. The device monitors and quantifies cell proliferation, morphology change, and provides an opportunity to observe the quality of cell attachment real-time. HSC-4 cells (1.0 x 10⁴) were seeded onto wells with varying concentrations of neem extract (25-800µg/ml) for varying times (0-48hrs). The cells were kept under observation, and IC₅₀ values were calculated for all extracts and those with DMSO only used as vehicle control. Neem leaf extract showed the best dose response (Fig. S1) and therefore chosen for further study. All subsequent experiments were carried out with the IC₅₀ concentration of neem extract unless otherwise specified.

Whole genome gene expression assays and data analyses

HSC-4 cells were seeded and treated with neem leaf extract (200µg/ml) at nine pre-defined time points (5min, 10min, 15min, 30min, 1hr, 3hr, 6hr, 10hr, and 24hr) and rescued with fresh

complete medium post 48-hour treatment at these time-points. The samples were assayed for gene expression using Illumina whole-genome HumanHT-12 v4 expression BeadChip (Illumina, San Diego, CA), following the manufacturer's specifications. RNA quality was checked using Agilent Bioanalyzer with RNA Nano6000 chip for integrity before being used in the gene expression assays. The RNA samples were labeled using Illumina TotalPrep RNA Amplification kit (Ambion, USA) and processed following the manufacturer's recommendations. Gene expression data was collected using Illumina's HiScan and analyzed with the GenomeStudio (v2011.1 Gene Expression module 1.9.0), and all assay controls were checked to ensure the quality of the assay and chip scanning. Raw signal intensities were exported from GenomeStudio for pre-processing and analyzed using R further. Gene-wise expression intensities from GenomeStudio were transformed, normalized using VST (Variance Stabilizing Transformation) and LOESS methods, respectively, using the R package lumi (Du et al., 2008) and further batch-corrected using ComBat (Johnson et al., 2007). The pre-processed intensities were subjected to differential expression analyses using the R package, limma (Ritchie et al., 2015) and fold change values were computed for various time-points in the treatment and rescue conditions versus the control samples. Genes that were serially up-regulated or serially down-regulated across at least four consecutive time-points were short-listed for further analysis.

Genome-wide shRNA lentivirus library screening and transduction

We used a genome-wide shRNA library for signaling pathway genes consisting of 27,500 shRNAs (Human Module 1, Cat No DHPAC-M1-P from Cellecta) for our experiment (Figure 1). Plating of the cells, transfection, DNase treatment, viral titer estimation, lentiviral transduction with polybrene and final puromycin selection were carried out following the manufacturer's instructions. HSC-4 cells were infected with the shRNA lentivirus library and cultured with or without neem extract treatment and harvested, and frozen, followed by the extraction of their genomic DNA (Figure 1). Two concentrations of neem leaf extract (200 µg/ml and 300 µg/ml) were used for the

treatment of cells. The RFP signals at the time of harvesting (72 hr after the lentivirus infection and before the puromycin treatment) were confirmed by observing the cells under a fluorescent microscope (Figure S2).

Next-generation sequencing post-genome-wide shRNA screening

We performed PCR amplification of barcode regions of the lentivirus library, followed by sequencing using the PGM system (Ion Torrent, Thermo Fisher, USA). Read QC was performed to meet the quality criteria for the reads in both control and shRNA library treated samples. The shRNAs with increased (> 2-fold) clone size in neem extract-treated cells compared to the control cells (with DMSO vehicle control) implied that the corresponding shRNAs contributed to the favorable cell survival or growth when treated with neem extract (Figure 2). On the other hand, the shRNAs whose clone sizes were decreased (< 2-fold) in neem extract-treated cells compared to the control cells implied that the knockdown of the corresponding genes worked in synergy with neem leaf extract for unfavorable cell survival or growth (Figure 2). Genes represented by at least two shRNAs were considered for further analyses to secure the reproducibility of the phenotypes of shRNA knockdown.

Interaction network, drug targets and pathway enrichment analyses

The short-listed genes from the whole genome expression assay and the consensus list of genes derived from the neem-treated shRNA transfected cells were combined after removal of redundancy if any. These genes were subjected to pathway enrichment analyses using two independent resources: namely, Speed (Parikh et al., 2010) and Reactome (Croft et al., 2014). Speed enables discovery of upstream over-represented signaling pathways responsible for observed differential expression changes in the genes, and Reactome allows merging of pathway identification and over-representation while allowing interactors to increase the analysis

background. We ran Speed while exercising the ‘Signature genes must be unique’ option to identify the significantly enriched pathway(s) first, before running it without choosing this option. This provided us with a complete list of genes perturbed in the enriched pathway(s). The perturbed list of genes from Speed and the over-represented genes and their interactors from enriched pathways detected by Reactome were subjected to network analyses using cBioPortal (Gao et al., 2013) (www.cbioportal.org) along with 50 significant most affected genomic profiles (mutations, putative copy number alterations from GISTIC and differentially expressed mRNAs (more than 2-fold) in the TCGA HNSCC dataset (provisional dataset)). Reactome was run by exercising the ‘Include Interactors’ option to increase the background list of genes. Cancer drugs were mapped to this network while differentiating between the FDA approved ones from those in the pipeline.

Results

IC₅₀ calculation of the neem extracts on HNSCC cell line HSC-4

The IC₅₀ values for the leaf, bark, fruit, seed and twig extracts were 0.184, 0.226, 2.932, 0.530, and 0.283 mg/ml, respectively, and the leaf extract showed the best dose response (Figure S1). Therefore, the leaf extract was used in subsequent experiments at IC₅₀ concentrations unless otherwise specified.

Gene expression changes in response to neem extract

The overall approach using a combination of functional genomics screening assay with the genome-wide shRNA library and the whole genome expression assay with treatment and rescue conditions is depicted in Figure 1. To precisely clarify the mode of molecular actions of the neem extract in a time-course manner, we investigated the changes in mRNA expression levels of all genes during the neem treatment compared to the steady state (24hr time-course), as well as their behaviors during the following rescue periods (24hr time-course).

Expression values for all genes at various time points (5min, 10min, 15min, 30min, 1hr, 3hr, 6hr, 10hr, and 24hr) and the follow-up rescue experiment (for the same periods of time) along with the fold change statistics and pathways affected is provided in Table S1. Expression of 28 genes was altered compared to no treatment, and their expression did not change significantly from the last time-point (24 hr) of neem leaf treatment (standard deviation across the last treatment time-point and in all rescue time-points ≤ 0.01). Out of those, 8 genes were temporally up-regulated (*CYP46A1*, *PDGFD*, *HOXA3*, *GRIK1*, *TRIM10*, *DDA1*, *OAZ1* and *NAT8B*) ($R \geq 0.6$), and 20 temporally down-regulated (*KLK12*, *PQLC1*, *LEF1*, *IL15RA*, *CHAC2*, *LMNA*, *EEF1A1*, *EHD4*, *CS*, *COG5*, *MS4A6A*, *FH*, *PRPS1*, *GR14*, *HNRNPM*, *OPRM1*, *KCNIP4*, *KCNC4*, *IMPACT* and *MRAP*) ($R \leq -0.6$). We selected genes ($n = 98$, 93 up-regulated and 5 down-regulated) with altered expression (ratio at least below 0.8 or at least above 1.2) over at least 4 consecutive time-points (Table S2) for pathway analysis comprising 15 cancer-related pathways (Table S1) for both treatment and rescue conditions (Figure 3). Genes involved in JAK-STAT, mTOR, VEGF, NF- κ B, and WNT signaling pathways, displayed a cumulative up-regulation effect over time, the action of neem extract was reversed upon the rescue. Their maximum values for expression change (\log_2 fc) were 0.75, 0.9, 0.6, 0.75 and 0.4, respectively. Genes involved in the ECM-receptor interaction, on the other hand, were cumulatively downregulated, that was completely reversed upon the rescue (Figure 3).

Functional genomics with pooled shRNA library with neem extract treatment

A representative result of the shRNA screening is shown in Figure 2. Sequencing read statistics, and unique shRNA reads obtained after next-generation sequencing post neem extract treatment are provided in Table 1. The results from the PCR amplification of the barcodes, in control and treated cells, showed 1369 and 1011 shRNAs with high (ratio versus control: >2) and with low (ratio versus control: <0.5) clonal sizes respectively (Figure 2). This represented 149 and

216 78 genes, respectively, where more than one shRNA were involved (Table S4). The direction of
217 changes in clonal size for genes are described in Table S4. To globally obtain a set of genes whose
218 function and/or expression were somehow related to the neem extract treatment, we investigated the
219 gene lists from the two independent screening experiments.

220

221 *Interaction network of genes and pathways picked from the gene expression and shRNA screening*
222 *assays*

223 We combined the short-listed genes whose expression were significantly changed by the
224 neem extract over at least 4 consecutive time-points (Table S2) and those derived from the
225 functional shRNA screening, resulting in a consolidated list of 321 non-redundant genes (Table S5).
226 These genes were considered to be affected somehow by the neem extract treatment either or both
227 in the functional and/or expression aspects; thus, we further used this list for pathway and
228 interaction network analyses to identify any molecular pathways influenced by the neem extract.

229 The list of 321 genes (Table S5) compiled from the combination of the serially up- or down-
230 regulated genes across at least four consecutive time-points (3hr, 6hr, 10hr and 24hrs) with large
231 expression change (ratio at least below 0.8 or at least above 1.2) and genes with significant changes
232 in the clonal sizes of shRNA-infected cells in the neem extract treatment, was subjected to pathway
233 enrichment analyses, using Speed and Reactome analysis tools independently. The enrichment
234 analyses using Speed while investigating only the unique set of signature genes resulted in only one
235 pathway, the TGF- β signaling pathway ($p = 7.67e-06$; FDR = $2.87E-05$), with a significant FDR
236 adjusted P value at the 95% level of confidence (Figure 4A). The unique genes perturbed in this
237 pathway were *TGIF1*, *SPHK1*, *DDIT3*, *RGS16*, *LRRC15*, *VDR* and *GABARAPL1*. When all genes
238 in the pathways were investigated, all signalling pathways passed the FDR significance threshold,
239 with MAPK_PI3K signalling pathway at the top, with the least FDR at $3.53e-13$ and 18 genes
240 perturbed out of 118 background genes, closely followed by TGF- β at an FDR of $4.69e-13$ and 19

genes perturbed out of 142 background genes (Figure 4B). The pathway enrichment analyses using Reactome resulted in three pathways with significant FDR (< 0.05): Attenuation phase (FDR = $3.56e-05$), *HSF-1*-dependent transactivation (FDR = $4.4e-05$) and *HSF-1* activation (FDR = 0.013) (Table S6). The genes over-represented in the Attenuation Phase pathway network were a subset of the two other *HSF-1*-related pathways; therefore, the *HSF-1*-related pathway was considered to be the enriched pathway in relation to the neem extract treatment. The perturbed genes common to both *HSF-1*-related pathways were *DNAJB1*, *HSPAIL*, *HSPA1A*, and *DEDD2*. *HSPB8* was unique to the *HSF-1*-dependent transactivation pathway network, while *RPA1* and *RPA2* were unique to the *HSF-1* activation pathway. *BAG3*, *PLK1*, and *SPHK1* were the interactors used in the analysis in the case of *HSF-1* activation pathway (Table S6).

Identifying potential drug targets for neem extract

Finally, we explored any candidate molecular targets that may work in synergy with the neem extract for HNSCC patients. Depending on the list of genes in the pathways which were considered to be related to the mode of actions of neem extract (Figure 4), we obtained the network and drug target interaction information for the 19 perturbed genes (*LRRC15*, *CITED2*, *RHOB*, *TIMP3*, *GADD45B*, *TGIF1*, *PIM1*, *SPHK1*, *PTGS2*, *DDIT3*, *DUSP1*, *DDIT4*, *MIR21*, *RGS16*, *HMOX1*, *HBEGF*, *EMP1*, *VDR* and *GABARAPL1*) from the TGF- β signalling pathway as detected by Speed (Parikh et al., 2010), the 10 genes (*DNAJB1*, *HSPAIL*, *HSPA1A*, *DEDD2*, *HSPB8*, *RPA1*, *RPA2*, *BAG3*, *PLK1* and *SPHK1*) from the *HSF-1*-based pathways as detected by Reactome, (Croft et al., 2014) and the 50 most altered genes for the TCGA HNSCC provisional data from the cBioPortal (Gao et al., 2013) (www.cbioportal.org). A network was drawn using cBioPortal contained 55 nodes, including 16 out of 28 query genes (Figure 5). Thirty-one percent of the network was categorized as ‘drug-target interactions’. Among these networks, *PTGS2*, *VDR*, *RHOB*, *PIM1*, *HSPAIL*, *HSPA1A*, and *SPHK1* were used as targets for cancer drug development,

previously. Out of these, *RHOB*, *VDR*, and *PTGS2* were the only candidates with FDA-approved anti-cancer drugs, with one each for *RHOB* and *VDR*, and 20 for *PTGS2* (Table S7). For others, quercetin (targeting *PIMI*) underwent seven clinical trials; and apricoxib (targeting *PTGS2*) underwent five clinical trials. Other genes such as *HMOX1* were the targets of 8 non-anti-cancer non-FDA-approved drugs (Table S7).

Discussion

Natural products are an important source of anti-cancer compounds (Hollosoy and Keri, 2004). Neem leaf extract and derived metabolites are shown to act as anti-cancer compounds specifically through the modulation of expression of critical cancer-related signalling molecules, including in oral cancer animal models (Harish Kumar et al., 2009; Paul et al., 2011; Hsieh et al., 2015; Kasala et al., 2015; Patel et al., 2016; Kowshik et al., 2017). Additionally, it is shown to act through modulation of the proinflammatory microenvironment in colorectal cancer (Gupta et al., 2013). Purified neem compounds have been shown to induce cytotoxicity and apoptosis, and affect cell cycle in various cancer cell lines (Roy et al., 2007; Priyadarsini et al., 2010; Chien et al., 2017), decrease tumor incidence and pre-neoplastic lesions in live hamster models of oral oncogenesis (Harish Kumar et al., 2010). This study aimed to identify the molecular function of the neem extract with a particular focus on its anti-HNSCC property.

A critical application of the RNA interference-based functional screening is the identification of drug mechanisms as well as the synergic drug target(s); on the other hand, global time-course gene expression profiling under drug treatment also provides us significant insights into the drug mechanisms. In the current study, we combined these two independent global experiments to investigate the molecular mechanisms of the neem extract with focuses on the anti-cancer properties for HNSCC.

291 Pathway enrichment analyses were conducted for shortlisted genes which were considered
 292 to be differentially affected somehow by the neem extract, either or both in functional and/or gene
 293 expression aspects, identifying a possible role for the TGF- β signaling pathway and the *HSF-1*
 294 activation network. The *HSF-1* is related to the TGF- β signaling pathway and connects the IGF,
 295 TGF- β and cGMP pathways, and controls development processes (Barna et al., 2012). The precise
 296 molecular interactions and biochemical downstream effects of the neem extract in combination with
 297 these pathways remain elusive and will be under a scope in the next research plan. We further raised
 298 a list of potential anti-HNSCC drug targets among genes in the TGF- β signaling pathway and/or
 299 the *HSF-1* activation network (Table S7) that may work in synergy with the neem extract against
 300 HNSCC. Although the potential effects of these drugs in synergy with the neem extract for HNSCC
 301 were not explored in this study, it is highly warranted to confirm their potential applicability in the
 302 clinical settings for HNSCC patients.

303 Despite the power of the pooled shRNA library screening, there have been caveats to the
 304 shRNA-based drug target discovery, and our study is no exception. Regarding the individual
 305 shRNA behavior in the shRNA screening, one would realize some biological discrepancies in
 306 comparison with the global gene expression and integrated pathway analyses. For example,
 307 although our pathway analyses identified the possible synergic interaction between neem extract
 308 and the TGF- β pathway, results from some of the individual hits in the shRNA screening were not
 309 concurrent. In the shRNA screening, knockdown of *SPHK1* and *LRRC15* (both are TGF- β
 310 pathway-related genes), for instance, were found to show no and positive effects on the cell growths
 311 of HSC-4 cells, respectively, under treatment of neem extract (Table S4). It can be hypothesized
 312 that the biological effects of the shRNA-mediated inhibition of single gene are heavily influenced
 313 by numerous experimental conditions: such as the time-scale (24hr in gene expression and six days
 314 in shRNA screening) and chemical activity (long-term culture in shRNA screening may inactivate
 315 the extract). One must also consider that single gene inhibition within a global web of signaling

networks may result in various levels of different biological homeostasis in the cells, resulting in cellular phenotypes different from those expected from the function of the pathway itself. Also, any single gene has multiple of down-stream effectors, possibly influencing both activation and inhibition of cell growth simultaneously. Despite such limitations in our shRNA screening, we identified potential candidate pathways by utilizing an integrated gene set obtained from either or both of the gene expression and/or shRNA screening. This signifies the importance of combining multiple of global datasets and screening hits to identify candidate pathways of drug compounds with broad spectrums of biological effects such as the neem extract.

In summary, our study points out that neem leaf extract may offer a therapeutic potential to treat patients with HNSCC through its synergic modulations of pathways via *HSF-1* activation and TGF- β signaling. These finding will hopefully help developing novel anti-tumor drugs to be used with the neem leaf extract against HNSCC.

Acknowledgements

The research is funded under the India-Japan Cooperative Science Programme by the Department of Science and Technology (DST), India to BP (Ref No. DST/INT/JSPS/P-181/2014), a grant from the Japanese Society of Promotion of Science (JSPS) to SI and funding from the Medical Research Institute, Tokyo Medical and Dental University for International Joint Research Project 2017. We thank PG Bharath Krishna for technical assistance.

Author Contributions: VP, HK, MP and RS: data generation; NMK and HK: data analysis, NMK, HK, BP and SI: manuscript writing, BP and SI: Overall supervision and study guidance.

Conflict of Interest: None

References

- Barna, J., Princz, A., Kosztelnik, M., Hargitai, B., Takacs-Vellai, K., and Vellai, T. (2012). Heat shock factor-1 intertwines insulin/IGF-1, TGF-beta and cGMP signaling to control development and aging. *BMC Dev Biol* 12, 32. doi: 10.1186/1471-213X-12-32.
- Blakely, K., Ketela, T., and Moffat, J. (2011). Pooled lentiviral shRNA screening for functional genomics in mammalian cells. *Methods Mol Biol* 781, 161-182. doi: 10.1007/978-1-61779-276-2_9.
- Brenner, J.C., Graham, M.P., Kumar, B., Saunders, L.M., Kupfer, R., Lyons, R.H., et al. (2010). Genotyping of 73 UM-SCC head and neck squamous cell carcinoma cell lines. *Head Neck* 32(4), 417-426. doi: 10.1002/hed.21198.
- Chien, S.Y., Hsu, C.H., Lin, C.C., Chuang, Y.C., Lo, Y.S., Hsi, Y.T., et al. (2017). Nimbolide induces apoptosis in human nasopharyngeal cancer cells. *Environ Toxicol* 32(8), 2085-2092. doi: 10.1002/tox.22423.
- Croft, D., Mundo, A.F., Haw, R., Milacic, M., Weiser, J., Wu, G., et al. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Res* 42(Database issue), D472-477. doi: 10.1093/nar/gkt1102.
- Du, P., Kibbe, W.A., and Lin, S.M. (2008). lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24(13), 1547-1548. doi: 10.1093/bioinformatics/btn224.
- Erler, J.T., and Linding, R. (2010). Network-based drugs and biomarkers. *J Pathol* 220(2), 290-296. doi: 10.1002/path.2646.
- Erler, J.T., and Linding, R. (2012). Network medicine strikes a blow against breast cancer. *Cell* 149(4), 731-733. doi: 10.1016/j.cell.2012.04.014.
- Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 6(269), p11. doi: 10.1126/scisignal.2004088.
- Gupta, S.C., Prasad, S., Sethumadhavan, D.R., Nair, M.S., Mo, Y.Y., and Aggarwal, B.B. (2013). Nimbolide, a limonoid triterpene, inhibits growth of human colorectal cancer xenografts by suppressing the proinflammatory microenvironment. *Clin Cancer Res* 19(16), 4465-4476. doi: 10.1158/1078-0432.CCR-13-0080.
- Harish Kumar, G., Chandra Mohan, K.V., Jagannadha Rao, A., and Nagini, S. (2009). Nimbolide a limonoid from *Azadirachta indica* inhibits proliferation and induces apoptosis of human choriocarcinoma (BeWo) cells. *Invest New Drugs* 27(3), 246-252. doi: 10.1007/s10637-008-9170-z.
- Harish Kumar, G., Vidya Priyadarsini, R., Vinothini, G., Vidjaya Letchoumy, P., and Nagini, S. (2010). The neem limonoids azadirachtin and nimbolide inhibit cell proliferation and induce apoptosis in an animal model of oral oncogenesis. *Invest New Drugs* 28(4), 392-401. doi: 10.1007/s10637-009-9263-3.
- Hollosy, F., and Keri, G. (2004). Plant-derived protein tyrosine kinase inhibitors as anticancer agents. *Curr Med Chem Anticancer Agents* 4(2), 173-197.
- Hsieh, Y.H., Lee, C.H., Chen, H.Y., Hsieh, S.C., Lin, C.L., and Tsai, J.P. (2015). Induction of cell cycle arrest, DNA damage, and apoptosis by nimbolide in human renal cell carcinoma cells. *Tumour Biol* 36(10), 7539-7547. doi: 10.1007/s13277-015-3477-0.
- Iorns, E., Ward, T.M., Dean, S., Jegg, A., Thomas, D., Murugaesu, N., et al. (2012). Whole genome in vivo RNAi screening identifies the leukemia inhibitory factor receptor as a novel breast tumor suppressor. *Breast Cancer Res Treat* 135(1), 79-91. doi: 10.1007/s10549-012-2068-7.
- Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA: a cancer journal for clinicians* 61(2), 69-90.

- Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1), 118-127. doi: 10.1093/biostatistics/kxj037.
- Kasala, E.R., Bodduluru, L.N., and Barua, C.C. (2015). Nimbolide inhibits invasion of breast cancer. *Cell Prolif* 48(2), 117-118. doi: 10.1111/cpr.12170.
- Kowshik, J., Mishra, R., Sophia, J., Rautray, S., Anbarasu, K., Reddy, G.D., et al. (2017). Nimbolide upregulates RECK by targeting miR-21 and HIF-1alpha in cell lines and in a hamster oral carcinogenesis model. *Sci Rep* 7(1), 2045. doi: 10.1038/s41598-017-01960-5.
- Mendes-Pereira, A.M., Sims, D., Dexter, T., Fenwick, K., Assiotis, I., Kozarewa, I., et al. (2012). Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. *Proc Natl Acad Sci U S A* 109(8), 2730-2735. doi: 10.1073/pnas.1018872108.
- Mishra, A., and Meherotra, R. (2014). Head and neck cancer: global burden and regional trends in India. *Asian Pac J Cancer Prev* 15, 537-550.
- Momose, F., Araida, T., Negishi, A., Ichijo, H., Shioda, S., and Sasaki, S. (1989). Variant sublines with different metastatic potentials selected in nude mice from human oral squamous cell carcinomas. *J Oral Pathol Med* 18(7), 391-395.
- Parikh, J.R., Klinger, B., Xia, Y., Marto, J.A., and Bluthgen, N. (2010). Discovering causal signaling pathways through gene-expression patterns. *Nucleic Acids Res* 38(Web Server issue), W109-117. doi: 10.1093/nar/gkq424.
- Patel, S.M., Nagulapalli Venkata, K.C., Bhattacharyya, P., Sethi, G., and Bishayee, A. (2016). Potential of neem (*Azadirachta indica* L.) for prevention and treatment of oncologic diseases. *Semin Cancer Biol* 40-41, 100-115. doi: 10.1016/j.semcancer.2016.03.002.
- Paul, R., Prasad, M., and Sah, N.K. (2011). Anticancer biology of *Azadirachta indica* L (neem): a mini review. *Cancer Biol Ther* 12(6), 467-476. doi: 10.4161/cbt.12.6.16850.
- Poell, J.B., van Haastert, R.J., Cerisoli, F., Bolijn, A.S., Timmer, L.M., Diosdado-Calvo, B., et al. (2011). Functional microRNA screening using a comprehensive lentiviral human microRNA expression library. *BMC Genomics* 12, 546. doi: 10.1186/1471-2164-12-546.
- Poell, J.B., van Haastert, R.J., de Gunst, T., Schultz, I.J., Gommans, W.M., Verheul, M., et al. (2012). A functional screen identifies specific microRNAs capable of inhibiting human melanoma cell viability. *PLoS One* 7(8), e43569. doi: 10.1371/journal.pone.0043569.
- Priyadarsini, R.V., Murugan, R.S., Sripriya, P., Karunakaran, D., and Nagini, S. (2010). The neem limonoids azadirachtin and nimbolide induce cell cycle arrest and mitochondria-mediated apoptosis in human cervical cancer (HeLa) cells. *Free Radic Res* 44(6), 624-634. doi: 10.3109/10715761003692503.
- Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7), e47. doi: 10.1093/nar/gkv007.
- Roy, M.K., Kobori, M., Takenaka, M., Nakahara, K., Shinmoto, H., Isobe, S., et al. (2007). Antiproliferative effect on human cancer cell lines after treatment with nimbolide extracted from an edible part of the neem tree (*Azadirachta indica*). *Phytother Res* 21(3), 245-250. doi: 10.1002/ptr.2058.
- Singleton, K.R., Kim, J., Hinz, T.K., Marek, L.A., Casas-Selves, M., Hatheway, C., et al. (2013). A receptor tyrosine kinase network composed of fibroblast growth factor receptors, epidermal growth factor receptor, v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, and hepatocyte growth factor receptor drives growth and survival of head and neck squamous carcinoma cell lines. *Mol Pharmacol* 83(4), 882-893. doi: 10.1124/mol.112.084111.
- Subapriya, R., and Nagini, S. (2005). Medicinal properties of neem leaves: a review. *Curr Med Chem Anticancer Agents* 5(2), 149-146.
- Tang, W., Dodge, M., Gundapaneni, D., Michnoff, C., Roth, M., and Lum, L. (2008). A genome-wide RNAi screen for Wnt/beta-catenin pathway components identifies unexpected roles for

438 TCF transcription factors in cancer. *Proc Natl Acad Sci U S A* 105(28), 9697-9702. doi:
 439 10.1073/pnas.0804709105.
 440 White, J.S., Weissfeld, J.L., Ragin, C.C., Rossie, K.M., Martin, C.L., Shuster, M., et al. (2007). The
 441 influence of clinical and demographic risk factors on the establishment of head and neck
 442 squamous cell carcinoma cell lines. *Oral oncology* 43(7), 701-712.
 443
 444

Table 1. Sequencing read statistics in functional pooled shRNA screening experiments.

	Total number of reads	Useful reads	%	Number of shRNA unique hits (total 27,495 shRNA)	%
DMSO control for neem leaf extract	7,658,211	6,762,717	88.3	27,260	99.1
Neem leaf extract (200µg/ml)	6,727,689	6,000,047	89.2	27,344	99.5
Neem leaf extract (300µg/ml)	7,620,025	6,621,474	86.9	27,397	99.6

448 **Figure Legends**

449 **Figure 1.** Overview of the screening method. A. Whole-genome functional shRNA screening
450 experiment, B. Whole-genome gene expression microarray experiment.

451 **Figure 2.** Scatter plot for barcodes in the sequencing reads in treated vs control neem leaf extract.

452 **Figure 3.** Cancer-related genes and pathways affected by neem extract at various time points and
453 with the follow up rescue.

454 **Figure 4.** Enriched pathways and over-represented genes analyzed by Speed using the options: A.
455 Unique genes and B. All genes.

456 **Figure 5.** Hypothetical gene-drug interaction network in relation to the effect of neem extract in
457 cancer.

459 **Supporting Figure and Table legends**

460 **Figure S1.** IC₅₀ of neem extracts on HSC-4 cells.

461 **Figure S2.** Cells after transduction (A) and the RFP signals at the time of harvesting (72 hr after the
462 lentivirus infection) by using a fluorescent microscope (B).

463 **Table S1.** Expression values for all the genes at various time points and the follow-up rescue
464 experiment along with the fold change statistics and cancer-specific pathways.

465 **Table S2.** Genes with altered expression across four or more time points.

466 **Table S3.** Results from the PCR amplification of the barcodes, in the control and in the leaf extract
467 treated samples and the ratios representing clonal size changes.

468 **Table S4.** Changes in clonal size from shRNA experiments with leaf extracts for drug target
469 analyses.

470 **Table S5.** Short- listed genes from the whole-genome gene expression (WGGX) assay and shRNA
471 experiments for pathway analyses.

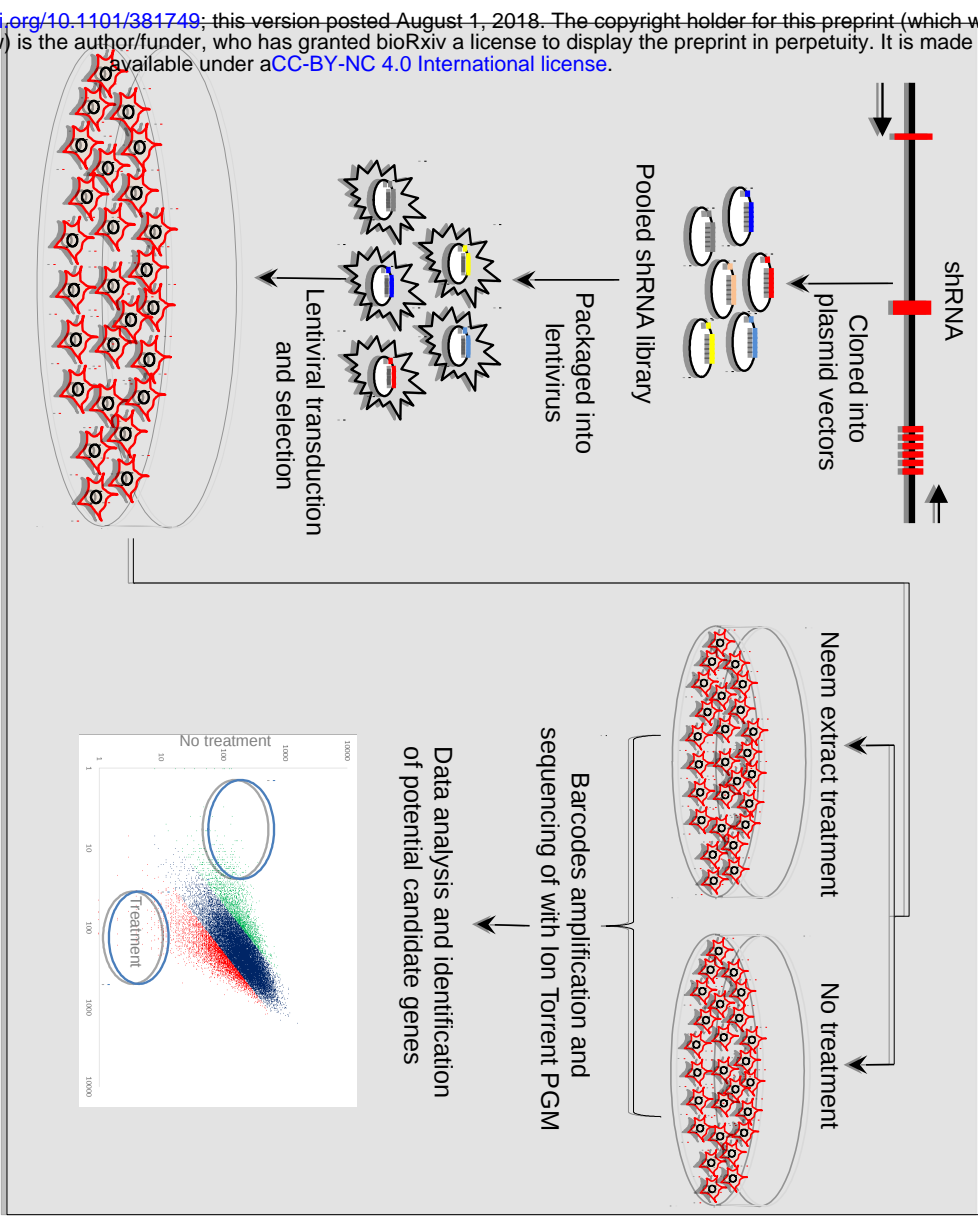
472 **Table S6.** Results from the pathway enrichment analyses performed using Reactome with a
 473 consensus list of 321 genes.

474 **Table S7.** Drug interactions with the genes perturbed in the *HSF-1*-based and the TGF- β signaling
 475 pathways.

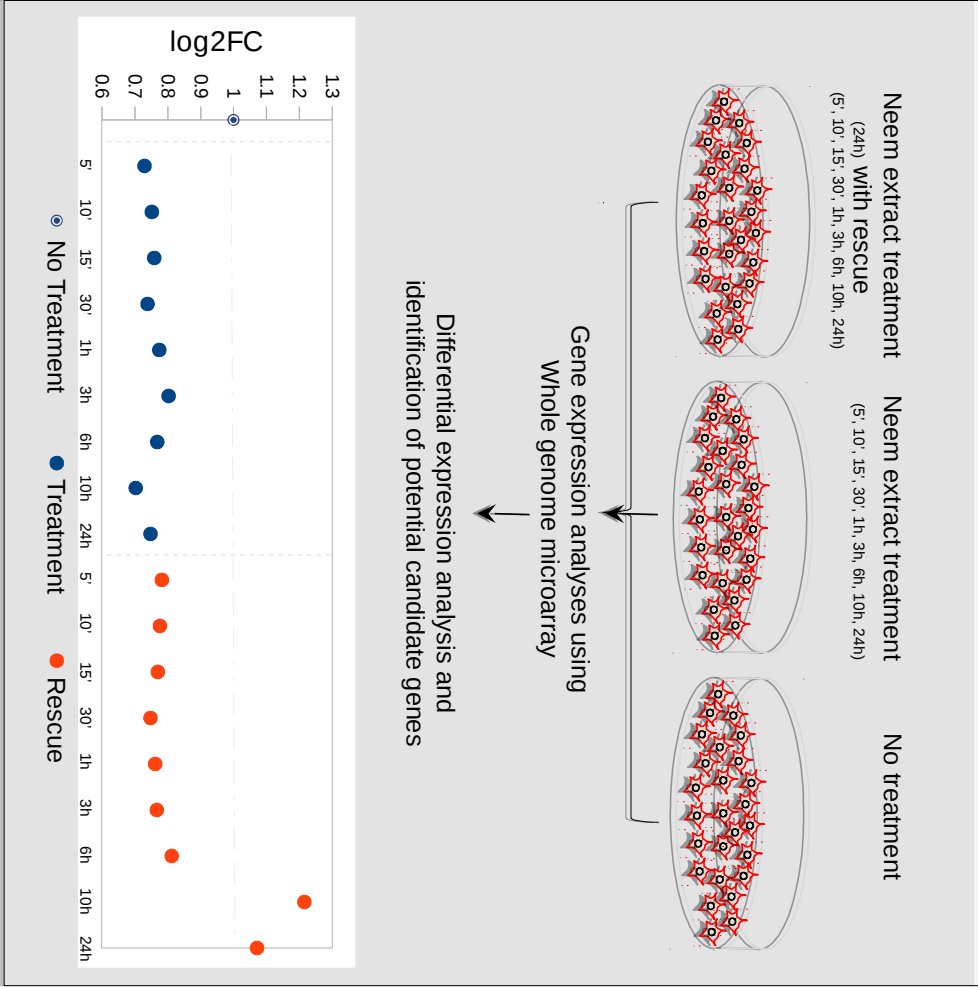
476

477

Genomics screen with pooled shRNA library

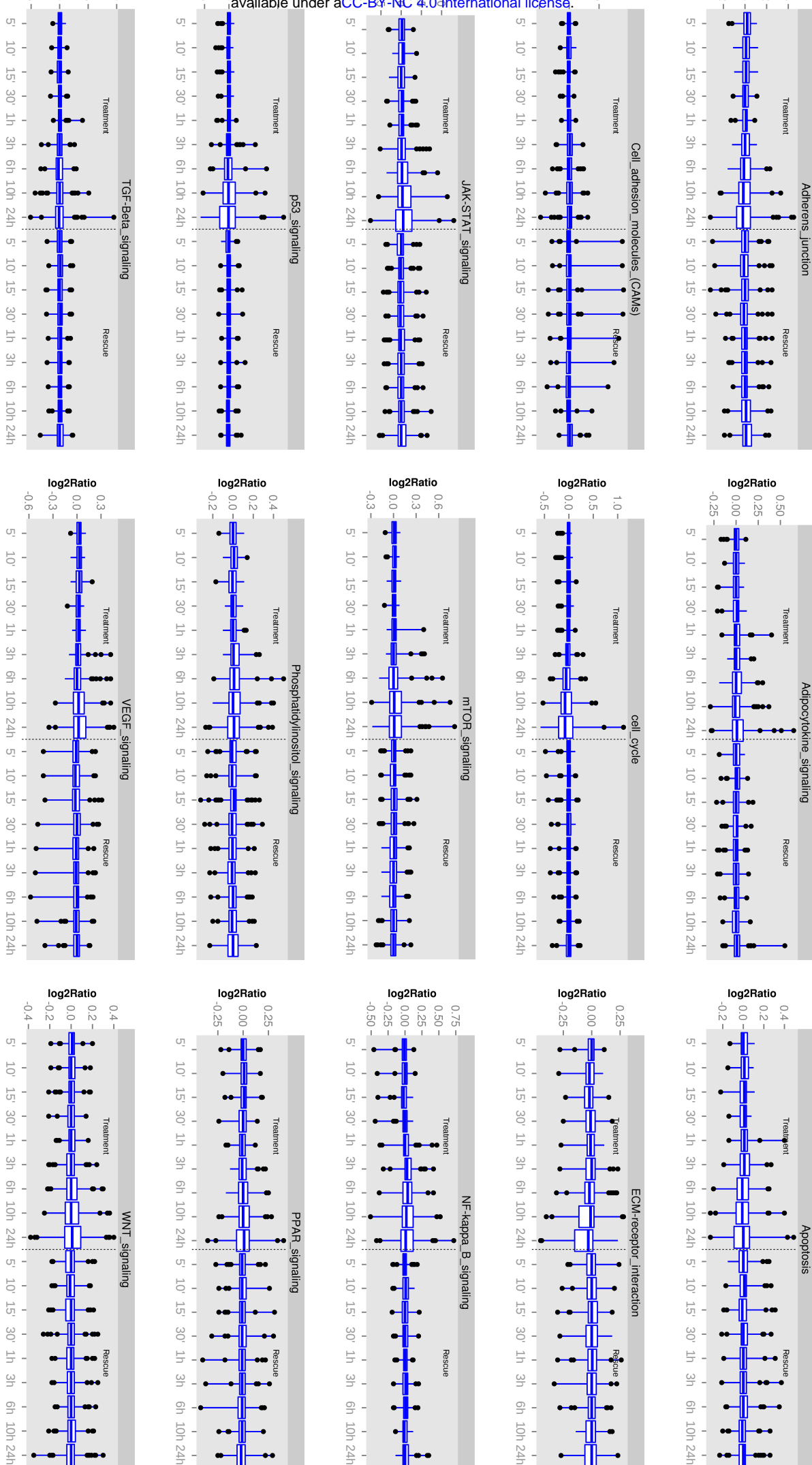


Whole genome microarray



Pathway enrichment analyses
using Speed and Reacome
Identification of drug targets
for enriched pathways

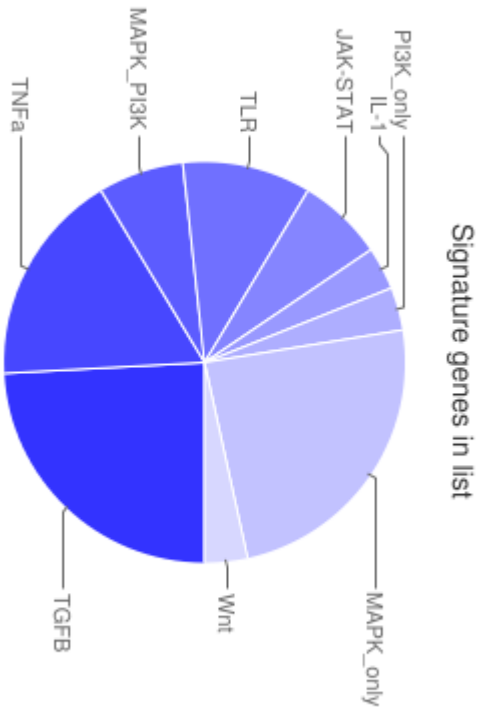
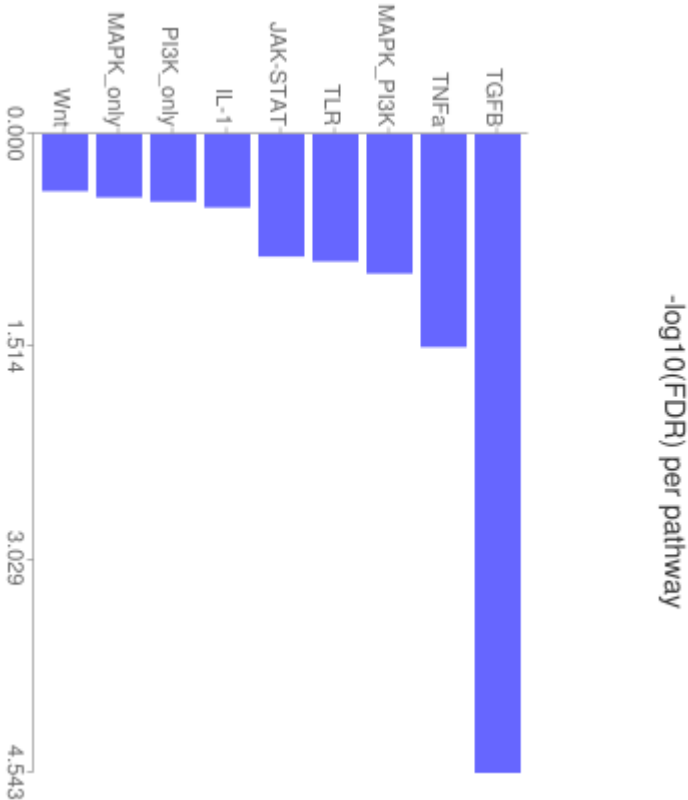




A. Unique genes

Total genes in list: 307
Total genes in background: 21485

Pathway	Genes_in_List	Genes_in_Background	P-Value	FDR	Genes
TGFB	7	51	7.67e-06	2.87e-05	TGIF1,SPHK1,DDIT3,RGS16,LRRC15,VDR,GABARAPL1 SPSB1,GFP12,SAMD4A,LITAF,TRRAP DNAJB1,CTGF DNAJB4,PEL1,ABL2 PKM2,ACTB LTB SPRED2 CCNE1,TAGLN,RGS2,HSPB8,GGCX,SFRP1,NR3C1 DNAJB9
TNFa	5	101	0.015	0.0306	
MAPK_P13K	2	24	0.0457	0.102	
TLR	3	73	0.0868	0.124	
JAK-STAT	2	48	0.15	0.135	
IL-1	1	23	0.282	0.301	
PI3K_only	1	27	0.322	0.331	
MAPK_only	7	459	0.484	0.355	
Wnt	1	49	0.506	0.393	
H2O2	0	23	1	1	

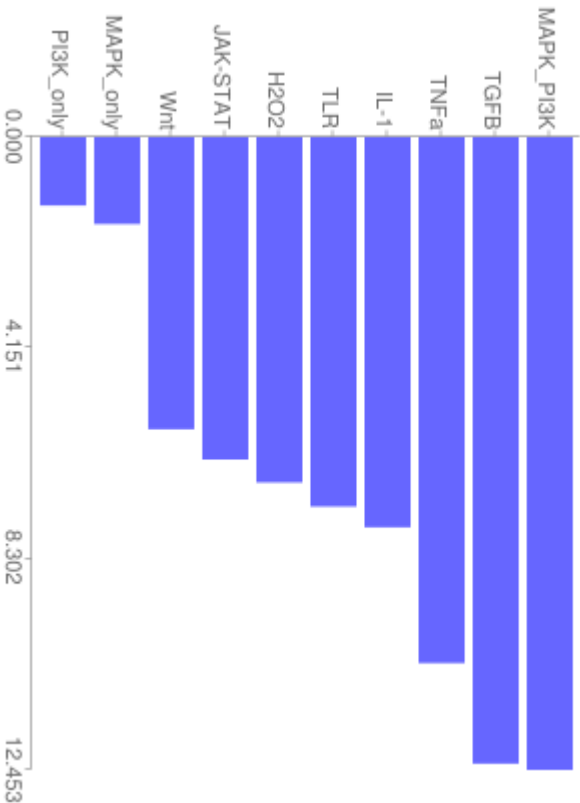


B. All genes

Total genes in list: 307
Total genes in background: 21485

Pathway	Genes in List	Genes in Background	P-Value	FDR
MAPK_Pi3K	18	118	8.11e-14	3.53e-13
TGFB	19	142	1.89e-13	4.69e-13
TNFa	22	259	2.54e-11	4.46e-11
IL-1	14	141	1.56e-08	2.1e-08
TLR	15	181	5.52e-08	5.37e-08
H2O2	9	60	1.72e-07	1.6e-07
JAK-STAT	11	114	7.39e-07	4.53e-07
Wnt	9	83	2.87e-06	1.79e-06
MAPK_only	14	559	0.0309	0.0195
PI3K_only	3	67	0.071	0.0452

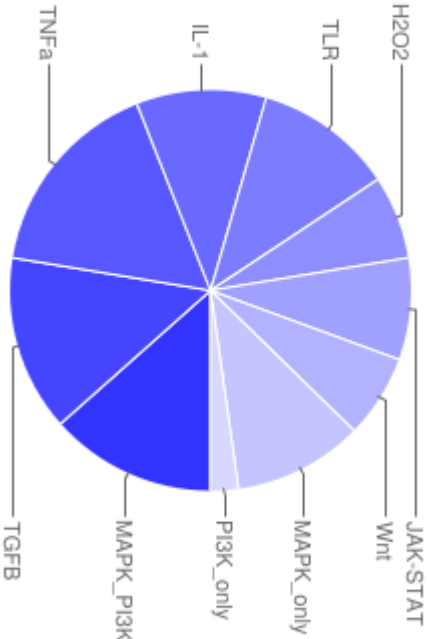
-log10(FDR) per pathway

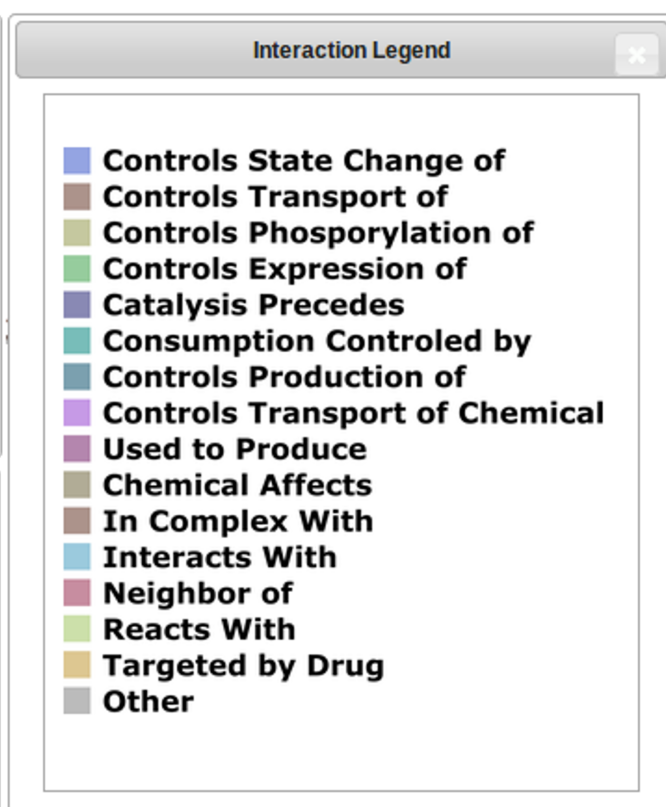
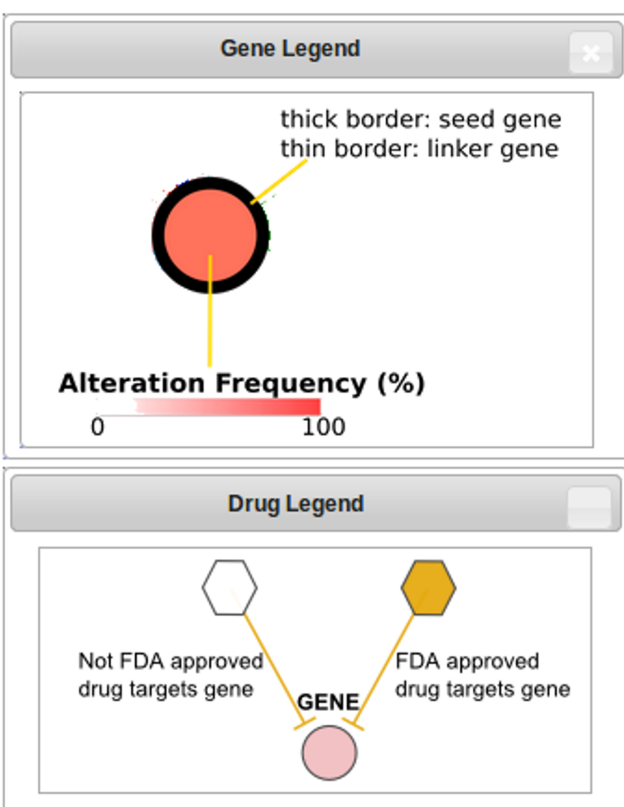
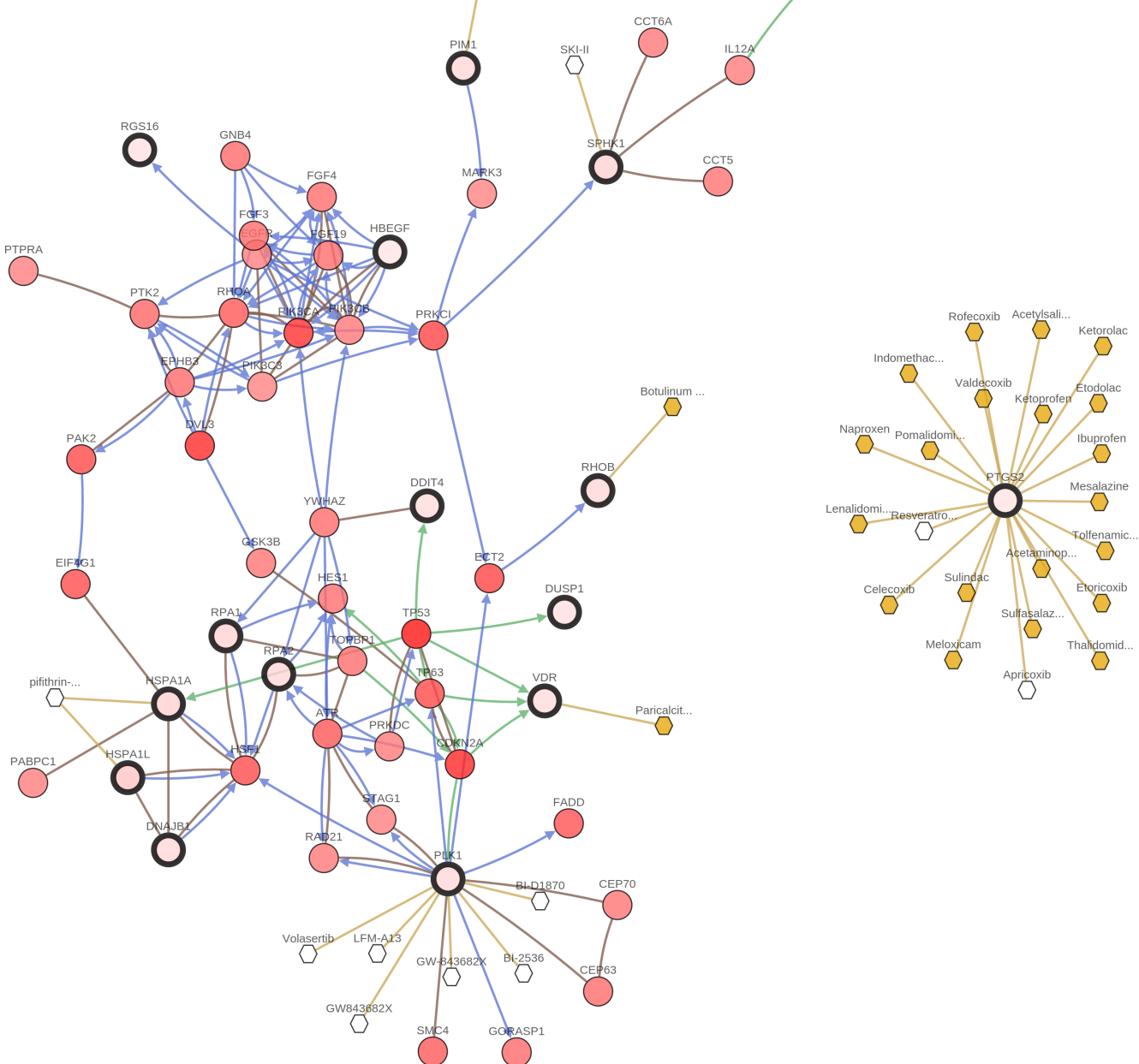


Genes

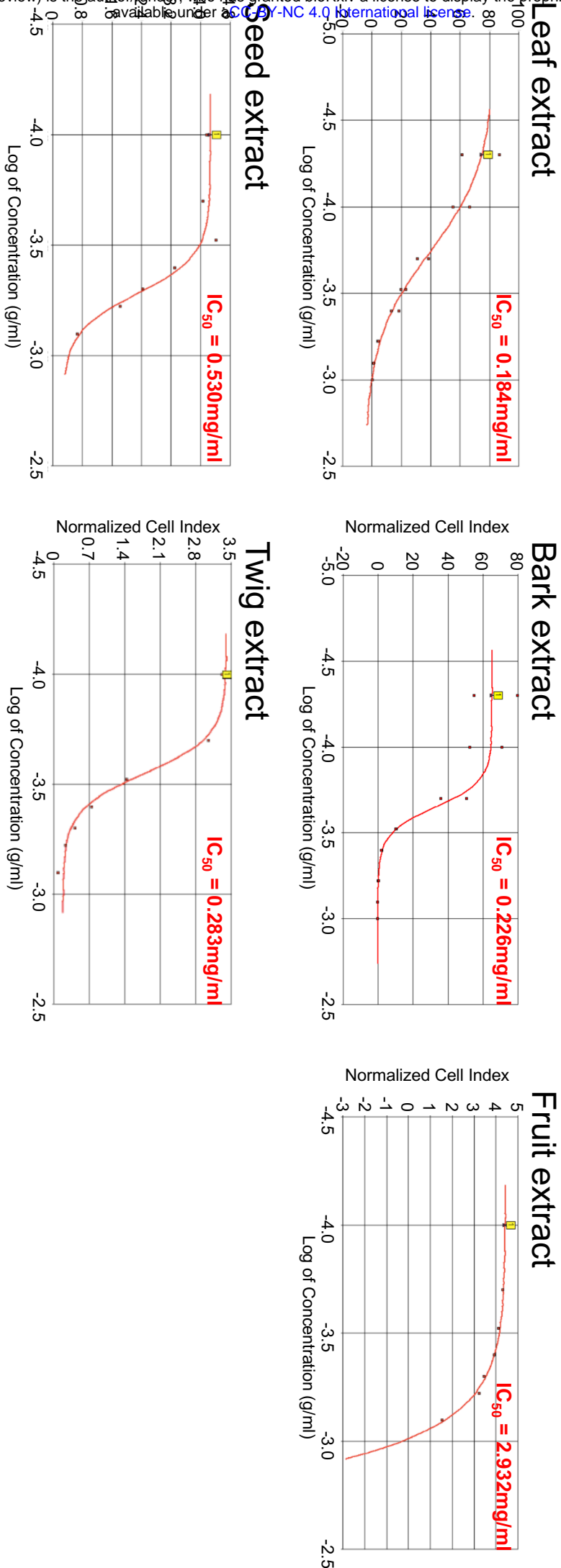
IL6,RHOB,ATF3,DNAJB1,KLF2,GEM,GDF15,PTGS2,CTGF,FOS,FOSB,DUSP1,MAFF,PMAIP1, WEE1,DDIT4,HBEGF,PPP1R15A
LRRC15,CITED2,RHOB,TIMP3,GADD45B,TGIF1,PIM1,SPHK1,PTGS2,DDIT3,DUSP1, DDIT4,MIR21,RGS16,HMOX1,HBEGF,EMP1,VDR,GABARAPL1
RHOB,GADD45B,LITAF,DDIT4,ZC3H12A,WEE1,PIM1,GDF15,SPSB1,FOS,DUSP1,ATF3, MAFF,GFP1T2,PPP1R15A,TRRAP,GEM,PTGS2,IL6,RELB,PMAIP1,SAMD4A
ATF3,GADD45B,DDIT4,GDF15,PTGS2,IL6,LTB,DUSP1,MAFF,ZC3H12A,PMAIP1,RELB,HBEGF, PPP1R15A
IL6,GADD45B,PELI1,SOC1,GEM,ATF3,PTGS2,DNAJB4,DUSP1,MAFF,ZC3H12A,PMAIP1,RELB, ABL2,PPP1R15A
RHOB,ATF3,GDF15,IL6,DUSP1,MAFF,PMAIP1,HMOX1,KLF2
PKM2,IL6,GADD45B,SOC1,PIM1,GEM,DDIT4,FOS,DUSP1,PMAIP1,ACTB
CITED2,RHOB,DDIT4,ATF3,FOS,FOSB,DUSP1,PMAIP1,DNAJB9
CCNE1,PMAIP1,TIMP3,IL6,GADD45B,NR3C1,EMP1,RGS2,GDF15,MIR21,HSPB8,GGCX,SFRP1, TAGLN
MAFF,DDIT4,SPRED2

Signature genes in list



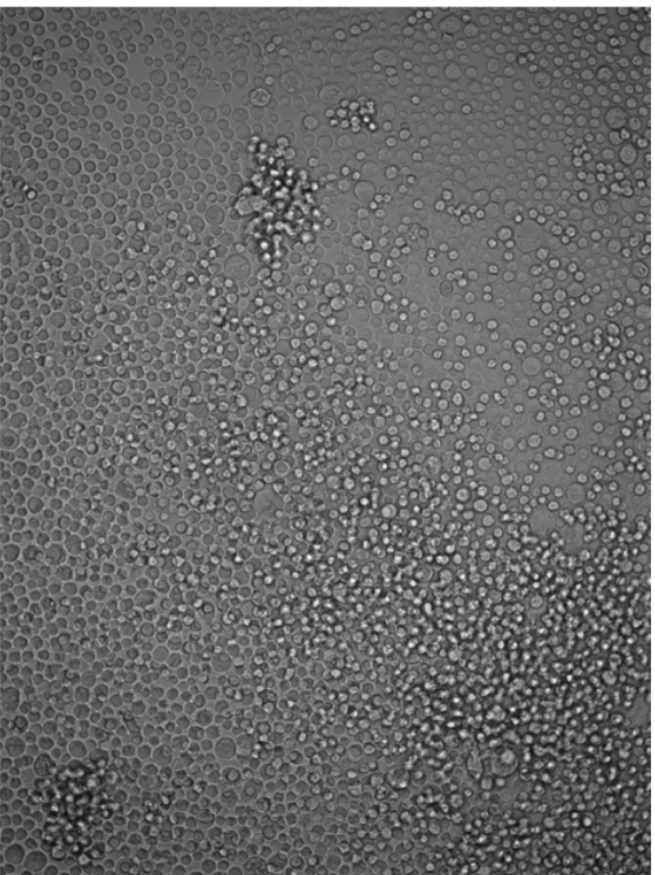


Supplementary Figure S1. IC₅₀ calculation using the xCelligence Real-time Cell Analysis (RTCA) DP instrument with various neem extracts and nimboide on HSC-4 cells.

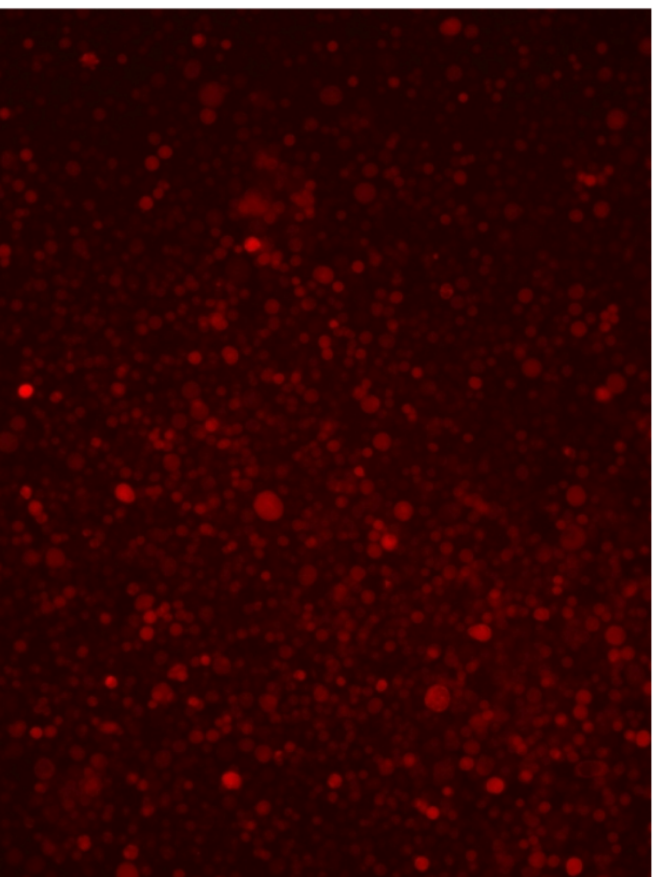


Supplementary Figure S2. HSC-4 Cells after lentivirus infection (A) and the RFP signals (B) detected by fluorescent microscope.

A



B



Supplementary Tables are too big and could not be uploaded here but are available upon request.