ORIGINAL ARTICLE

Comparative Copy Number Variation Profiling of GL01, an Immortalized Non-small Cell Lung Cancer Cell Line Derived from a Filipino Patient, and A549 Lung Adenocarcinoma Cells

Treena Rica D. Teh,[1] Kim Claudette J. Fernandez,[1] Maria Katrina Diana M. Cruz, RCh, MSc,[1]
Patrick Gabriel G. Moreno, MSc,[1] Ruel C. Nacario, PhD,[2]
Gladys C. Completo, PhD[2] and Francisco M. Heralde, III, MS, PhD[1,3]

[1]*Department of Biochemistry and Molecular Biology, College of Medicine, University of the Philippines Manila, Manila, Philippines*
[2]*Institute of Chemistry, University of the Philippines Los Baños, Los Baños, Laguna, Philippines*
[3]*Molecular Diagnostics and Cellular Therapeutics Laboratory, Department of Pathology, Lung Center of the Philippines, Quezon City, Philippines*

ABSTRACT

Background and Objectives. Cell lines serve as invaluable tools in studying lung cancer biology and developing new therapies to combat the disease. However, commercially available cell lines are typically of Caucasian origin and may be less representative of the local genetic background. To address this, our lab previously immortalized cells from pleural fluid of a Filipino non-small cell lung cancer (NSCLC) patient via CDK4 transduction. Copy number variations (CNVs) are a type of genetic variation which may affect physiology and disease by disrupting gene function or altering gene expression, and in cancer, these may be associated with patient outcomes. CNV profiling can be valuable for understanding the biology of our immortalized cells and identifying genes that could serve as potential targets for diagnostic, prognostic, and therapeutic interventions. This study aimed to characterize previously immortalized NSCLC-derived cells, GL01, in comparison with an established lung adenocarcinoma (LUAD) cell line, A549, through whole-genome microarray-based copy number profiling.

Methods. DNA was extracted from GL01 and A549 cells using a commercially-available silica-based DNA extraction kit. DNA extracts were quantified and normalized for microarray analysis. Whole-genome copy number profiling was done using the OncoScan CNV Plus Assay following the manufacturer's protocols, and data was analyzed using the Chromosome Analysis Suite software. Functional analysis of genes identified to be involved in copy number aberrations was done using the PANTHER Classification System.

Results. Copy number aberrations span 1,592,737,105 bp in GL01 and 1,715,708,552 bp in A549, with a high degree of concordance between the two. Large-scale and focal copy number aberrations previously identified to be recurrent in various LUAD cohorts were present in both GL01 and A549. Focal copy number aberrations associated with previously described lung cancer-related genes involve the PDE4D gene in GL01 and the SKIL and CDKN2A/CDKN2B genes in both GL01 and A549. PANTHER Pathway analysis of genes positively correlated with mRNA expression showed that the ubiquitin proteasome pathway was significantly overrepresented in both GL01 (FDR p = 0.000074) and A549 (FDR p = 0.000075), with 20 genes involved. Additionally, the KRAS:p.G12C/S:c.34G>T/A somatic mutation variant was detected in both GL01 and A549.

Conclusion. This study provides a method for identifying potentially clinically-relevant genes associated with a

Poster presentation – International Conference of the Korean Society of Molecular and Cell Biology, November 6, 2023, ICC, Jeju, South Korea.

Corresponding author: Treena Rica D. Teh
Department of Biochemistry and Molecular Biology
College of Medicine
University of the Philippines Manila
547 Pedro Gil St. Ermita, Manila 1000, Philippines
Email: tdteh3@up.edu.ph

sample's copy number aberrations and the pathways they represent, providing personalized mechanistic, prognostic, and therapeutic insights into the cancer biology of our cells.

Keywords: copy number variation, copy number aberration, copy number alteration, non-small cell lung cancer, lung adenocarcinoma, OncoScan, PANTHER, somatic mutation

INTRODUCTION

Lung cancer remains to be one of the most common and deadly forms of cancer, causing over 1.7 million deaths worldwide in 2020.[1] Among the histological subtypes of lung cancer, non-small cell lung cancer (NSCLC) is the most common form, comprising about 85% of cases.[2] Cancer cell lines serve as invaluable tools in studying lung cancer biology and developing new therapies to combat the disease. However, commercially available cell lines, including those of lung cancer, are typically of Caucasian origin and may be less representative of the local genetic background, potentially resulting in phenotypic differences that may limit the generalizability of research using these cell models.[3]

To address the need for Filipino-derived lung cancer cell lines, our lab previously isolated and immortalized cells from a Filipino patient diagnosed with lung adenocarcinoma (LUAD).[4] To further our understanding on these cells as a potential Filipino patient-derived NSCLC cell line, it is important to characterize them. One way by which this can be done is through copy number variation (CNV) profiling.

CNVs are a type of genetic variation in which the number of copies of repeated segments of DNA, at least around 1 kb in size, vary between individuals of a population.[5,6] These variations may come in the form of gains or losses of copies with respect to an artificial reference genome, a parental genome, or the individual's germline genome, such as in the case of somatic variants.[5] While most CNVs do not result in significant phenotypic variability, they can potentially influence physiology and disease by disrupting gene function or affecting gene expression through gene dosage alterations.[7]

Cancer genomes feature extensive CNV, some of which have been associated with patient prognoses and treatment response in LUAD.[8] In a previous pan-cancer analysis, CNVs have been found to play a role in differential expression of genes.[9] CNVs have also been associated with lung cancer immune evasion,[10-12] which may impact disease course and response to immunotherapeutic modalities. Moreover, another pan-cancer study found that at least 50% of NSCLC cases are primarily driven by copy number alterations rather than point mutations.[13] These highlight the importance of looking into copy number alterations to better understand the biology of our immortalized cells and perhaps suggest genes that can serve as potential targets for diagnostic, prognostic, and therapeutic interventions down the line.

Thus, the main objective of this study is to characterize previously immortalized NSCLC-derived cells, GL01, in comparison with an established lung adenocarcinoma cell line, A549, through whole-genome microarray-based copy number profiling. Specifically, it aims to identify copy number aberrations (CNAs) present in the genomes of GL01 and A549 cells using the OncoScan CNV Plus Assay and to describe the functional significance of the identified aberrations through PANTHER Pathway analysis.

MATERIALS AND METHODS

Patient Recruitment and Sample Collection

The patient from which the cells were derived is a 58-year-old female diagnosed with moderately differentiated primary lung adenocarcinoma. Immunohistochemistry of the patient's primary tumor showed positive staining for thyroid transcription factor 1 (TTF1) and Napsin A. Pleural effusion was collected from the patient at the Molecular Diagnostics and Cellular Therapeutics Laboratory, Lung Center of the Philippines, in March 2019 as part of routine tumor sampling for genetic profiling.

Primary Cell Isolation and Immortalization

Pleural fluid was diluted with an equal volume of 1X PBS with ciprofloxacin (4 μg/mL) and centrifuged at 300 × g for 10 minutes at room temperature to separate cells from cell-free fluid. Cells were resuspended in red blood lysis buffer and incubated for five minutes at room temperature. The suspension was centrifuged, and the resulting pellet was resuspended in fresh Dulbecco's Modified Eagle Medium/ Nutrient Mixture F-12 (DMEM/F-12) + GlutaMAX (Gibco) supplemented with 5% non-heat inactivated fetal bovine serum (FBS) (Gibco), insulin (10 μg/mL) and ciprofloxacin (4 μg/mL) + 30% cell-free fluid. Cells were seeded in low-adhesion cell culture dishes at a cell density ≥1-1.5 x 10^6 live cells/mL. Cells were grown for three weeks with addition of 25% fresh medium every week. Resulting cells were collected, washed, and incubated with Accutase (Gibco) at 1.5 mL/dish for 5 min at 37 °C, 5% CO_2. Cells were then diluted in DMEM/F-12 + GlutaMAX (Gibco) supplemented with 5% non-heat inactivated FBS, insulin (20 μg/mL) and ciprofloxacin (4 μg/mL) to achieve a density of ≤0.5 x 10^6 live cells/mL. In culture, primary cells grew in monolayers and had epithelial morphology similar to A549 LUAD cells.[4]

Cells were seeded in a 6-well plate at 200,000 cells/well seeding density and grown for 24 hours to 50% confluency. Cells were then incubated with recombinant CDK4 lentivirus (Applied Biological Materials LV611) at a multiplicity of infection of 5 and polybrene (8 ug/mL) (Applied Biological Materials). After six hours incubation, the culture medium was replaced with fresh media and incubated for 72 hours. The cells were then subcultured to a tissue culture flask containing puromycin (2 μg/mL) (Applied Biological Materials) for

10 days. Colonies with puromycin resistance were selected and propagated in a puromycin-containing medium. Cells successfully transduced with CDK4 were labeled as GL01 and further propagated using standard cell culture propagation protocols. CDK4 gene expression was confirmed using reverse transcription quantitative polymerase chain reaction (RT-qPCR).[4]

Cell Culture

GL01 cells and established lung adenocarcinoma cells, A549 (ATCC CCL-185™), were cultured using standard cell culture propagation protocols. GL01 and A549 cells were cultured and maintained in RPMI-1640 Medium (Gibco) and DMEM (Gibco), respectively. Cultures were supplemented with 10% FBS (Gibco) and 1% Pen-Strep (Gibco) and were incubated at 37˚C with 5% CO_2. After the fourth passage, a portion of cells were cryopreserved through suspension in 10% dimethyl sulfoxide (DMSO) (Sigma-Aldrich) as backup for subsequent analysis.

DNA Extraction and Normalization

DNA was extracted from cryopreserved cells using DNeasy Blood & Tissue Kits (QIAGEN) following the manufacturer's protocols. DNA was quantified using a Quantus™ Fluorometer (Promega) and normalized to 12 ng/uL using 1X Tris-EDTA (TE) buffer for subsequent analysis.

Whole-Genome Copy Number Profiling

Whole-genome copy number profiling was done using the OncoScan CNV Plus Assay Kit (Applied Biosystems) following the manufacturer's protocol. Briefly, DNA sample was annealed with the CNV and Somatic Mutation Probe Mix overnight at 58˚C. Following the hybridization, the mix was divided into two for filling gaps in the circularized probe with A/T and G/C nucleotides. An exonuclease specific for linear DNA, incomplete circular probe and excess probes was used to select gap-filled probes. The circular probe was then linearized by endonuclease cleavage, biotinylated, and amplified. The amplified fragments were subjected to *HaeIII* endonuclease digestion and hybridized to the array overnight, stained, and scanned using the GeneChip® System (Affymetrix). Scanning results were analyzed with the Chromosome Analysis Suite (ChAS) version 4.4 software using default settings (Applied Biosystems). CN gains or losses determined in this study were hereinafter referred to as copy number aberrations (CNAs) rather than copy number variations (CNVs), which typically refers to germline variations that vary between individuals.

Functional Analysis of Genes

Functional analysis of genes involved in regions of CNA was performed through PANTHER (Protein Analysis Through Evolutionary Relationships) classification system (http://www.pantherdb.org). PANTHER integrates genomes, gene function classifications, pathways, and statistical analysis tools in order to investigate genome wide experimental data.[14] In this system, gene functions are classified through consideration of both published scientific evidence and evolutionary relationship. Analysis of data was accomplished with PANTHER version 17.0 (released 2022-02-22) as instructed by the developers.[14] For further analysis, genes were narrowed down to those which are correlated with mRNA expression based on publicly available LUAD datasets from The Cancer Genome Atlas (TCGA).[15] Additionally, genes involved in regions of CNA that are known to be cancer drivers, oncogenes, or tumor suppressors in LUAD were identified using CancerMine, a literature-mined database of drivers, oncogenes, and tumor suppressors in various cancers.[16]

Statistical Analysis

Statistical analysis of pathway overrepresentation was done using Fisher's Exact test with the Benjamini-Hochberg false discovery rate (FDR) correction for multiple testing, which is the default algorithm for PANTHER's overrepresentation test. A pathway was considered overrepresented when there are more genes involved in that particular pathway than expected, which is based on the proportion of genes in the whole human genome which are involved in that pathway.[14] Results were considered statistically significant at FDR $p < 0.05$.

Ethical Clearance

The study was approved by the Lung Center of the Philippines – Institutional Ethics Review Board (IERB No. LCP-CS-003-2018). Written informed consent for the primary cell culture and its subsequent analyses was obtained from the patient.

RESULTS

Microarray Quality Control (QC) Metrics

Percent aberrant cells determined using the ChAS software's TuScan algorithm showed that the GL01 sample was composed of 90% aberrant cells, while the A549 sample was homogenous. Ploidy state determined using the same algorithm showed that both samples had a ploidy of 3. To assess the internal validity of microarray results, the following QC metric thresholds, as recommended by the software developers, were used: MAPD ≤0.30, nd SNP QC ≥26.00, and nd waviness SD ≤0.12. Both GL01 and A549 samples were within bounds for these metrics (Table 1). Additionally, both samples were identified to have sufficient "normal diploid" markers for signal calibration (Low Diploid Flag = 0). Overall, these metrics show that the microarray results obtained in this study are valid for further analysis.

Whole-Genome Copy Number Profile of GL01 and A549

A total of 59 CNV segments spanning 1,592,737,105 bp were identified in GL01, while 103 CNV segments spanning 1,715,708,552 bp were identified in A549 (Figure

1, Supplementary Table S1). Overlapping regions of CNV in GL01 and A549 span a total size of 1,553,471,197 bp, with 98.84% of GL01's CNAs also found in A549, and 90.76% of A549's CNAs also found in GL01. This suggests a high degree of concordance between the copy number profiles of GL01 and A549.

Large-scale amplifications present in both GL01 and A549 previously identified in various LUAD cohorts include amplifications in 1q, 2p, 3q, 5p, 7p, 7q, 8q, 11q, 15q, 16p, 17q, 19q, 20p, 20q, and 21q.[17-23] Other large-scale amplifications found in GL01 and A549 which are not recurrently found in LUAD include amplifications in 3p, 8p, 9p, 9q, 10p, 10q, 11p, 12p, 12q, 14q, 16q, 17p, 23p, and 23q, which may be patient- or tumor-specific amplifications. No large-scale deletions were detected in GL01 and A549. Focal copy number losses previously reported in various LUAD cohorts found in our samples include a focal copy number loss in 9p21.3 of both GL01 and A549, and in 18q23 of A549.[19,24,25] A focal copy number gain in 3q26.2, previously observed in LUAD,[23] was also found in both GL01 and A549. Focal copy number gains previously reported in LUAD found within large-scale amplifications of both GL01 and A549 include 1q21.2, 2p15, 5p14.3, 5p15.31, 5p15.33, 7p11.2, 7q21.2, 8p11.23, 8q21.11, 8q21.13, 8q24.12, 8q24.21, 11q13.3, 12p12.1, 12p14.1, 12q15, 14q13.3, 17q12, 19q12, 19q13.12, and 20q13.32.[19,23-25]

Non-recurrent focal copy number gains found in our samples include a gain in 5q35.2-35.3 of both GL01 and A549; 5q11.2-12.1 of GL01; and 1p36.33-36.32, 6p21.31, and 18q23 of A549. On the other hand, non-recurrent focal copy number losses found in our samples include a loss in 3p14.2, 3q13.31, and 8p11.22 of both GL01 and A549; 6p21.33 of GL01; and 4p16.3, 5q12.1, 6p21.32-21.31, 6q26, and 6q27 of A549. Similar to non-recurrent large-scale amplifications, these may represent patient- or tumor-specific CNAs.

Genes Involved in CNAs of GL01 and A549

Regions of CNA in GL01 involve 16,305 genes, 204 of which are not involved in CNAs of A549. In A549, 17,406 genes are involved in CNAs, 1,305 of which are not found

Table 1. OncoScan Quality Control (QC) Metrics of GL01 and A549

Cells	Threshold Test	MAPD (≤0.30)	nd SNP QC (≥26.00)	nd waviness SD (≤0.12)	Low Diploid Flag	TuScan %AC	TuScan Ploidy
GL01	Within bounds	0.28	34.58	0.0878	No	90%	3
A549	Within bounds	0.29	34.36	0.1187	No	Homogenous	3

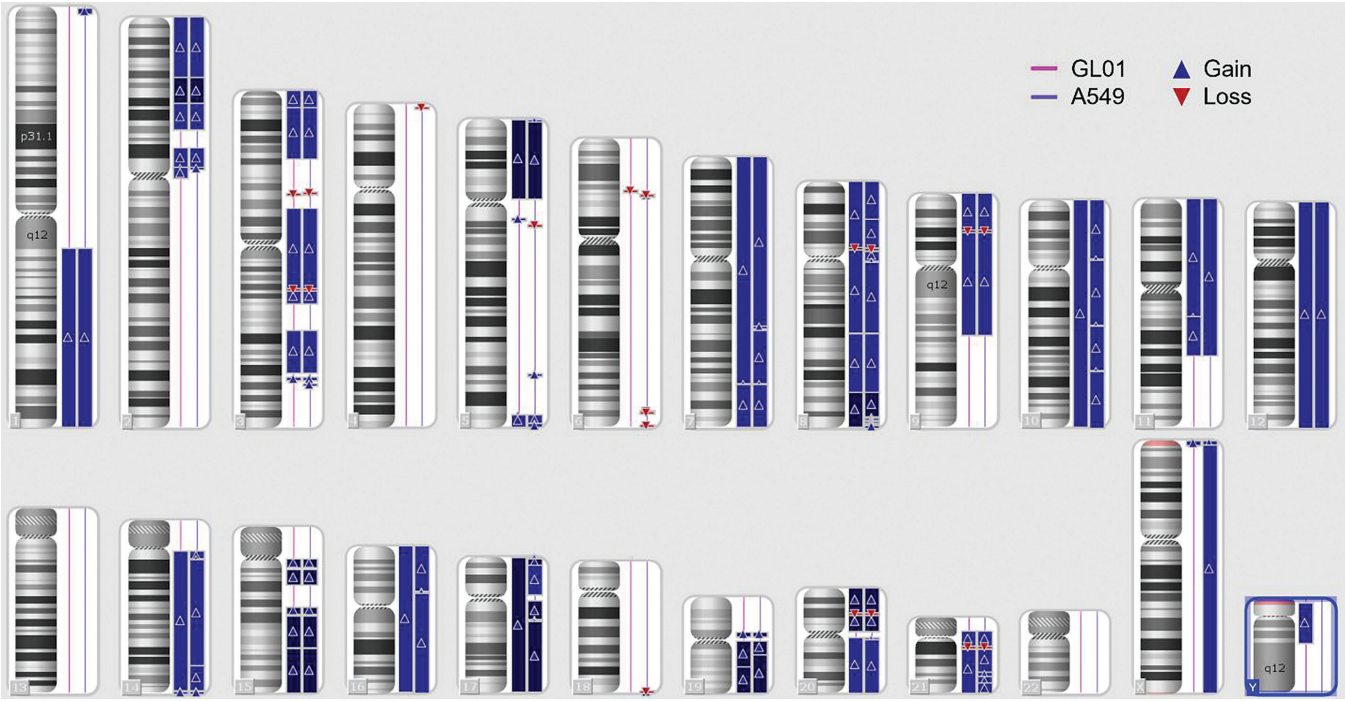

Figure 1. Global copy number analysis of GL01 and A549 cells. Each chromosome is represented in separate boxes, with copy number aberrations of GL01 and A549 shown beside each chromosome in pink and purple, respectively. Copy number gains are indicated as blue lines with arrowheads, while copy number losses are indicated as red lines with arrowheads. Detailed information on these segments can be found in Supplementary Table S1.

in CNAs of GL01. A total of 16,101 genes are involved in CNAs of both A549 and GL01. Among genes involved in CNAs of GL01, 317 genes were identified by CancerMine to be cancer drivers, oncogenes, or tumor suppressors in LUAD (Supplementary Table S2). In A549, 348 genes were identified, including all 317 genes identified in GL01 (Supplementary Table S2). Significant lung cancer-related genes found in recurrent focal CNAs of GL01 and A549 are listed in Table 2. Overall, these may serve as candidate genes contributing to CNA-mediated cancer progression in GL01 and A549 and warrant further investigation in our samples.

Pathway Analysis of Genes

PANTHER Pathway analysis of all genes involved in CNAs of both GL01 and A549 showed no statistically significant pathway overrepresentation. To make our data more relevant in the context of gene expression, we narrowed down the gene lists to those that are correlated with mRNA expression based on publicly available LUAD datasets from TCGA.[15] From this, 1,623 genes involved in CNAs of GL01 were identified to be correlated with mRNA expression, and these genes could be functionally classified into 100 PANTHER pathways. Pathways with at least two-fold overrepresentation in either GL01 or A549 mRNA expression-correlated genes involved in CNAs are shown in Figure 2. Among these pathways, only the ubiquitin proteasome pathway was found to be significantly overrepresented (FDR p = 0.000074), with 20 genes involved (Table 3). A parallel analysis in A549 identified 1,632 genes, likewise classifiable into 100 PANTHER pathways with significant overrepresentation of the ubiquitin proteasome pathway (FDR p = 0.000075), involving the same 20 genes (Table 3). These findings suggest that CNAs may have a possible role in driving this pathway in GL01 and A549.

Table 2. Focal Copy Number (CN) Aberrations in GL01 and A549 previously Identified to have Lung Cancer-related Genes[23]

Cells	Chromosome	Cytoband	CN	Type	Genes
GL01, A549	3	q26.2	3.00	Gain	SKIL
GL01	5	q11.2	2.50	Gain	PDE4D
GL01, A549	9	p21.3	0.00	Loss	CDKN2A/CDKN2B

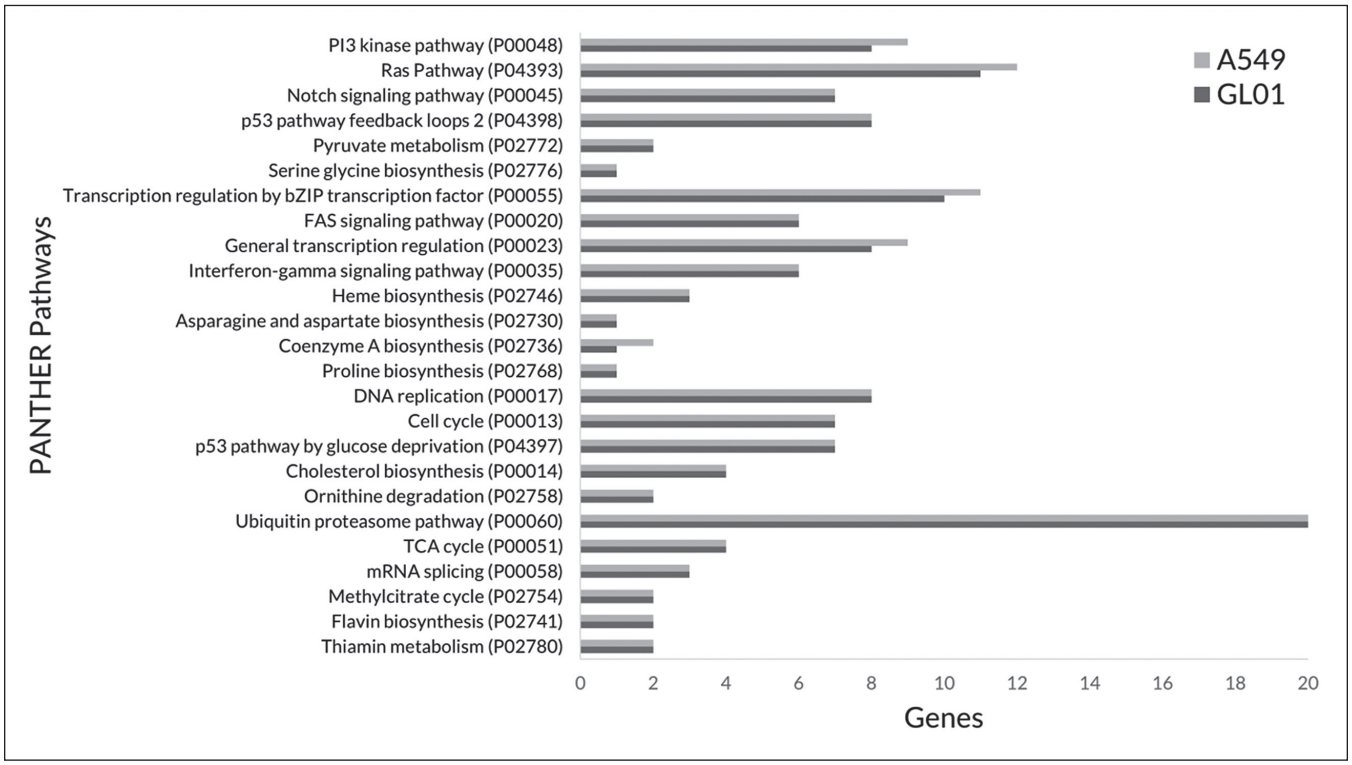

Figure 2. PANTHER pathways represented by mRNA expression-correlated genes involved in copy number aberrations of GL01 and A549. Pathways with at least two-fold overrepresentation in either GL01 or A549 are shown in this graph. Only the ubiquitin proteasome pathway was found to be significantly overrepresented in GL01 (FDR p = 0.000074) and A549 (FDR p = 0.000075).

Table 3. Ubiquitin Proteasome Pathway-related Genes which are Correlated with mRNA Expression and are Involved in Copy Number (CN) Aberrations of GL01 and A549

Gene	Name	CN GL01	CN A549	Type
PSMD3	26S proteasome non-ATPase regulatory subunit 3	4.00	4.00	Gain
UBE2G1	Ubiquitin-conjugating enzyme E2 G1	4.00	4.00	Gain
PSMC5	26S proteasome regulatory subunit 8	4.00	4.00	Gain
PSMC4	26S proteasome regulatory subunit 6B	5.00	4.00	Gain
PSMD4	26S proteasome non-ATPase regulatory subunit 4	3.00	3.00	Gain
WWP2	NEDD4-like E3 ubiquitin-protein ligase WWP2	3.00	3.00	Gain
PSMD11	26S proteasome non-ATPase regulatory subunit 11	4.00	4.00	Gain
PSMC2	26S proteasome regulatory subunit 7	3.00	3.00	Gain
MDM2	E3 ubiquitin-protein ligase Mdm2	3.00	3.00	Gain
UBE3B	Ubiquitin-protein ligase E3B	3.00	3.00	Gain
UCHL5	Ubiquitin carboxyl-terminal hydrolase isozyme L5	3.00	3.00	Gain
WWP1	NEDD4-like E3 ubiquitin-protein ligase WWP1	3.00	3.00	Gain
SMURF1	E3 ubiquitin-protein ligase SMURF1	3.00	3.00	Gain
PSMD9	26S proteasome non-ATPase regulatory subunit 9	3.00	3.00	Gain
UBA2	SUMO-activating enzyme subunit 2	5.00	5.00	Gain
PSMD7	26S proteasome non-ATPase regulatory subunit 7	3.00	3.00	Gain
ITCH	E3 ubiquitin-protein ligase Itchy homolog	3.00	3.00	Gain
UBE3C	Ubiquitin-protein ligase E3C	3.00	3.00	Gain
UBE2D4	Ubiquitin-conjugating enzyme E2 D4	3.00	3.00	Gain
UBE2G2	Ubiquitin-conjugating enzyme E2 G2	3.00	2.50	Gain

Table 4. Somatic Mutation Variants Detected with High Confidence in GL01 and A549 Cells Using OncoScan CNV Plus Assay

Cells	Somatic Mutation Variant	Mutation Score (≥6.00)	Chromosome	Position
GL01	KRAS:p.G12C/S:c.34G>T/A	7.82	12	25,245,351
A549	KRAS:p.G12C/S:c.34G>T/A	21.79	12	25,245,351

Detection of Somatic Mutations in GL01 and A549

Aside from copy number data, the OncoScan CNV Plus Assay is also able to detect somatic mutations in 64 mutation hotspots across 9 genes known in cancer. Among those, the KRAS:p.G12C/S:c.34G>T/A mutation was detected with high confidence in both GL01 and A549 (Table 4). Additionally, this mutation was found to be involved in a chromosome-wide copy number gain (CN = 3.00) observed in both GL01 and A549. These findings contribute to the possible mechanisms of cancer progression in GL01 and A549.

DISCUSSION

CNAs are emerging as a key mediator of lung cancer pathogenesis, with mounting evidence of its association with various clinical outcomes.[8] Thus, characterizing CNAs may provide insights into the underlying mechanisms of lung cancer biology, proposing potential diagnostic, prognostic, and therapeutic targets down the line.[26] To our knowledge, this is the first study looking into CNAs of NSCLC cells derived from a Filipino patient, particularly using the whole-genome microarray-based OncoScan assay. OncoScan characterizes whole-genome copy number profiles with adequate sensitivity, lower cost, consistency, and a streamlined, cancer-focused approach, providing a rapid, simple, and reproducible platform intended for clinical diagnostics.[27] This reduces the complexity of data interpretation, in contrast to the intensive data analyses needed to process the broad genomic information that sequencing-based approaches provide. The assay, however, is limited to characterizing CNAs and a panel of somatic mutation hotspots, unlike sequencing-based approaches that can detect other genetic variations such as single nucleotide variants, small insertions/deletions, and other less common or novel mutations.[28] Further characterization of GL01 through sequencing-based approaches is thus recommended to validate the results of this study and allow a more comprehensive analysis of its genomic landscape.

Large-scale CNAs involve thousands of genes and are important drivers of CNA-mediated changes in gene expression. In GL01 and A549, the observation of large-scale

CNAs which are recurrent in LUAD provide clues on possible chromosome- or arm-wide events that may be important in the development of LUAD. However, due to the large number of genes involved in large-scale CNAs, it is generally difficult to pinpoint genes that are of actual clinical relevance.[5] As was done in this study, using tools such as CancerMine may aid in prioritization of known cancer-related genes for further investigation, especially when dealing with genome-wide data that would otherwise require extensive literature review.[16] CancerMine was developed based on previous text-mining approaches used to extract gene-disease relations and has an average precision of 85.6% in classifying genes as a cancer driver, oncogene, or tumor suppressor.[16] In contrast to manually curated databases of cancer genes, text-mined databases such as CancerMine are advantageous in that they can provide results without the curator cost, time, and effort needed to keep a database up-to-date.[16] CancerMine has previously been utilized to characterize genome-wide driver gene, oncogene, and tumor suppressor gene mutational and transcriptional signatures in bulk and single-cell datasets across various cancers, including lung cancer,[29] chronic myeloid leukemia,[29] colorectal cancer,[30] and melanoma.[31] Such studies provide a catalogue of genes that may be crucial in the evolution of cancers as they grow and progress, allowing the identification of mechanisms that potentially targetable for diagnostic and therapeutic purposes.[6] Although text-mining tools are limited in that they cannot provide an exhaustive list of all known cancer-related genes, redundancy of research on these genes in literature allows such tools to achieve a better recall.[16] Given its technical limitations, the use of CancerMine may benefit from further literature and experimental validation, which can be a future direction for this study.

In contrast to large-scale CNAs, focal CNAs (≤3 Mb in size) contain fewer genes and likely persist through an evolutionary process as a tumor develops.[32] They are thus postulated to be more highly enriched with cancer driver genes that favor tumor progression, as observed previously.[33] Because of this, focal CNAs are ideal for identifying candidate genes that drive hallmarks of cancer in a particular tumor.

In the present study, known lung cancer-related genes involved in focal CNAs included SKIL, PDE4D, and CDKN2A/CDKN2B genes. The SKIL gene encodes a protein that inhibits the transforming growth factor-β (TGF-β)/Smad pathway, promoting cell proliferation by impeding growth inhibitory signals of TGF-β.[34] Expression of the SKIL gene has been found to be elevated in NSCLC and has been associated with a malignant phenotype and T cell immune escape.[35] Thus, the amplification of this gene in GL01 and A549 may contribute to their cancer characteristics. PDE4D is a gene encoding an isoform of phosphodiesterase-4, which has previously been demonstrated to promote proliferation and angiogenesis of lung cancer.[36] A copy number gain of this gene found only in GL01 may thus confer a unique survival advantage that may be worth further investigation in comparison with A549 that does not possess this copy number gain. Lastly, CDKN2A/CDKN2B genes encode cyclin dependent kinase inhibitors which act as tumor suppressors by regulating the cell cycle.[37] Hence, deletion of this gene results in uncontrolled proliferation of cells characteristic of cancer. True enough, a deletion of this gene has been correlated with poor survival outcomes in LUAD patients.[38] Altogether, these genes involved in focal CNAs provide prognostic insights about their tumors of origin and may serve as potential personalized therapeutic targets.

The ubiquitin proteasome pathway regulates protein degradation, a process that is crucial in maintaining protein homeostasis. Thus, alterations in the expression of components of this pathway, including ubiquitin, enzymes, ligases, and the 26S proteasome, may result in dysregulation of cellular protein levels.[39] In fact, dysregulations in the process of ubiquitination and deubiquitination, primarily through changes in the expression of deubiquitinating enzymes (DUBs), have been associated with lung cancer through their effects on cell cycle control and DNA damage response and repair mechanisms.[40-44] Thus, components of the ubiquitin proteasome system have attracted attention in recent decades as potential therapeutic targets for cancer. For instance, bortezomib, a proteasome inhibitor, has shown efficacy in the treatment of multiple myeloma and has been approved by the FDA for such use.[45]

In this study, 20 mRNA expression-correlated genes related to the ubiquitin proteasome pathway were found to be involved in CNAs of GL01 and A549, resulting in a statistical overrepresentation of the pathway. In the context of cancer, pertinent among these is MDM2, a ubiquitin ligase that regulates the degradation of the tumor suppressor p53. Elevations in MDM2, observed in many cancers, cause increased degradation of p53, leading to uncontrolled cell proliferation.[39] Various compounds targeting MDM2, such as Nutlin-3a, RG7388, RG7112, and ATSP-7041, are in development against various types of cancer, including lung cancer, neuroblastoma, retinoblastoma, leukemia, and melanoma.[39,46-49]

KRAS mutations are frequently observed in many cancers including LUAD, where it is estimated to occur in 20-25% of cases.[8] The G12C mutational variant, found in GL01 and A549, makes up the majority of KRAS-mutant LUAD cases.[50] A concomitant copy number gain in a mutated *KRAS* gene, also found in GL01 and A549, has been postulated to bring about synergistic effects to tumor aggressiveness and progression through its association with tumors that are larger and less differentiated.[51] Moreover, an amplified KRAS mutation has previously been associated with worse prognosis than either mutation or amplification alone.[52] While there is currently no established treatment for LUAD with KRAS mutations, $KRAS^{G12C}$-specific inhibitors, adagrasib and sotorsarib, are currently in phase 1 and 2 studies with promising results.[53,54]

A previous study on gene alterations across lung cancer cell lines have identified several recurrent features, including an inactivation at TP53, CDKN2A, or RB1,[55] consistent with our finding of a deletion of CDKN2A/CDKN2B in both GL01 and A549. The same study noted mutual exclusivity of mutations in genes belonging to the same biological pathway, such as KRAS and EGFR genes that participate in the same signal transduction pathway.[55] Consistent with this, a mutation in EGFR was not observed in both GL01 and A549, which already harbor a KRAS mutation. As was also observed by this study, multiple genetic alterations may be present and potentially cooperate in lung carcinogenesis of a single cell line.[55] Alterations which are more often observed across a larger set of samples make them more likely to be key drivers of lung carcinogenesis.[56] Thus, we hypothesize that the KRAS mutation and deletion of CDKN2A, both of which have been commonly observed in other lung cancer cell lines, may play an important role in the carcinogenesis of GL01 and A549. This, however, does not discount the potential impacts of the other CNA-associated genes identified in this study, as genetic heterogeneity across cell lines exist.[55,56] Experimental validation is warranted to confirm the biological significance of these gene alterations in GL01.

Overall, the findings of this study provide many insights into copy number-related characteristics of GL01 in comparison to an established cell line. However, it is important to note that not all genomic alterations translate to a phenotypic change, and that observations in this study must be validated experimentally to better understand the biology of GL01 in relation to CNAs. Moreover, the temporal relationship between the copy number aberrations observed in this study and the development of cancer in the original tumor cannot be established from the presented data. It is not known whether these variations are germline CNVs that resulted to cancer, or if these are CNAs that arose as the cancer progressed or as the cell lines were established and cultivated. The impact of the immortalization process on the CNAs observed in this study also cannot be ascertained due to a lack of valid CNA data on the primary cancer cells. Further study on the germline and primary cancer cell CNV profiles of the source patient would be valuable in determining the role of these CNAs in patient-specific cancer progression. Additionally, as this is a pioneering study on the CNV profile of only one Filipino patient-derived NSCLC cell line, CNAs and their associated genes and pathways identified in this study cannot be generalized to all NSCLC and LUAD cells. Development and characterization of more cell lines are necessary to establish hypotheses on copy number changes and their genotypic and phenotypic implications in lung cancer biology. Nevertheless, copy number profiles generated in this study serve as baseline information that can be used for future comparisons, leading to better insights about cell lines that can be used to better model lung cancer in the local setting.

CONCLUSION

GL01 and A549 share a similar copy number and somatic mutation profile, with large-scale amplifications and focal CNAs previously identified to be recurrent in various LUAD cohorts. Pertinent lung cancer-related genes associated with these CNAs include SKIL and CDKN2A/CDKN2B genes in both GL01 and A549, and PDE4D in GL01. CNAs may play a role in driving the ubiquitin proteasome pathway in GL01 and A549, among other identified CNA-associated genes and a somatic mutation in *KRAS* that potentially drive cancer in these cells. Observed alterations in CDKN2A and KRAS, which have previously been identified in other lung cancer cell lines, are hypothesized to be key drivers of carcinogenesis in GL01 and A549. This study demonstrated a means of identifying potentially clinically relevant genes involved in CNAs of a sample and what pathways may be represented by these genes, providing mechanistic and therapeutic insights into a novel Filipino-patient derived lung cancer cell line. However, experimental testing of identified genes or pathways is warranted, as not all CN changes represent a phenotypic change. Although observations are not generalizable to all NSCLC, data generated by this study may serve as baseline information for future comparisons in the local setting.

Acknowledgments

The authors would like to acknowledge the Molecular Diagnostics and Cellular Therapeutics Laboratory (MDCTL), Lung Center of the Philippines, and the Molecular Diagnostics and Multi-omics Laboratory (MDML), Department of Biochemistry and Molecular Biology, UP College of Medicine for the facility use and support, Dr. Jose B. Nevado Jr. for lending us equipment for microarray analysis, and Ms. Aurora S. Nakpil for providing us with some reagents.

Statement of Authorship

All authors certified fulfillment of ICMJE authorship criteria.

Author Disclosure

All authors declared no conflicts of interest.

Funding Source

This research is part of a project funded by the Philippine Commission on Higher Education (CHED) through the Philippine – California Advanced Research Institute (PCARI-IHITM 2017–18) grant given to Dr. Ruel Nacario and Dr. Gladys Completo of UPLB-IC.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021 May;71(3):209-49. doi: 10.3322/caac.21660 PMID: 33538338.
2. Schabath MB, Cote ML. Cancer progress and priorities: lung cancer. Cancer Epidemiol Biomarkers Prev. 2019 Oct;28(10):1563-79. doi: 10.1158/1055-9965.EPI-19-0221. PMID: 31575553; PMCID: PMC6777859.
3. Badal S, Campbell KS, Valentine H, Ragin C. The need for cell lines from diverse ethnic backgrounds for prostate cancer research. Nat Rev Urol. 2019 Dec;16(12):691-2. doi: 10.1038/s41585-019-0234-y. PMID: 31520083; PMCID: PMC7219314.
4. Fernandez KCJ, Moreno PGG, Macapagal DD, Cabanatan MM, Barzaga MTA, Tan-Liu N, et al. Immortalization of lung adenocarcinoma primary cells from a Filipino patient using CDK4 lentivirus. 2024. Unpublished.
5. Pös O, Radvanszky J, Buglyó G, Pös Z, Rusnakova D, Nagy B, et al. DNA copy number variation: main characteristics, evolutionary significance, and pathological aspects. Biomed J. 2021 Oct;44(5): 548-59. doi: 10.1016/j.bj.2021.02.003. PMID: 34649833; PMCID: PMC8640565.
6. Pös O, Radvanszky J, Styk J, Pös Z, Buglyó G, Kajsik M, et al. Copy number variation: methods and clinical applications. Appl Sci. 2021 Jan; 11(2):819. doi: 10.3390/app11020819.
7. Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet. 2009;10:451-81. doi: 10.1146/annurev.genom.9.081307.164217. PMID: 19715442; PMCID: PMC4472309.
8. Testa U, Pelosi E, Castelli G. Molecular characterization of lung adenocarcinoma combining whole exome sequencing, copy number analysis and gene expression profiling. Expert Rev Mol Diagn. 2022 Jan;22(1):77-100. doi: 10.1080/14737159.2022.2017774. PMID: 34894979.
9. Shao X, Lv N, Liao J, Long J, Xue R, Ai N, et al. Copy number variation is highly correlated with differential gene expression: a pan-cancer study. BMC Med Genet. 2019 Nov 9;20(1):175. doi: 10.1186/s12881-019-0909-5. PMID: 31706287; PMCID: PMC6842483.
10. Athie A, Marchese FP, González J, Lozano T, Raimondi I, Juvvuna PK, et al. Analysis of copy number alterations reveals the lncRNA ALAL-1 as a regulator of lung cancer immune evasion. J Cell Biol. 2020 Sep 7;219(9):e201908078. doi: 10.1083/jcb.201908078. PMID: 32858747; PMCID: PMC7480115.
11. Bhattacharya A, Bense RD, Urzúa-Traslaviña CG, de Vries EGE, van Vugt MATM, Fehrmann RSN. Transcriptional effects of copy number alterations in a large set of human cancers. Nat Commun. 2020 Feb 5;11(1):715. doi: 10.1038/s41467-020-14605-5. PMID: 32024838; PMCID: PMC7002723.
12. McGranahan N, Rosenthal R, Hiley CT, Rowan AJ, Watkins TBK, Wilson GA, et al; TRACERx Consortium. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. Cell. 2017 Nov 30;171(6):1259-1271.e11. doi: 10.1016/j.cell.2017.10.001. PMID: 29107330; PMCID: PMC5720478.
13. Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. Nat Genet. 2013 Oct;45(10):1127-33. doi: 10.1038/ng.2762. PMID: 24071851; PMCID: PMC4320046.
14. Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, et al. Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat Protoc. 2019 Mar;14(3):703-21. doi: 10.1038/s41596-019-0128-8. PMID: 30804569; PMCID: PMC6519457.
15. Broad Institute TCGA Genome Data Analysis Center. Correlations between copy number and mRNA expression [Internet]. 2016 Jan 28 [cited 2023 Jul 4]. Broad Institute of MIT and Harvard. Available from: https://gdac.broadinstitute.org/runs/analyses__2016_01_28/reports/cancer/LUAD-TP/Correlate_CopyNumber_vs_mRNA/nozzle.html Referenced in doi: 10.7908/C1DF6QM9
16. Lever J, Zhao EY, Grewal J, Jones MR, Jones SJM. CancerMine: a literature-mined resource for drivers, oncogenes and tumor suppressors in cancer. Nat Methods. 2019 Jun;16(6):505-7. doi: 10.1038/s41592-019-0422-y. PMID: 31110280.
17. Balsara BR, Sonoda G, du Manoir S, Siegfried JM, Gabrielson E, Testa JR. Comparative genomic hybridization analysis detects frequent, often high-level, overrepresentation of DNA sequences at 3q, 5p, 7p, and 8q in human non-small cell lung carcinomas. Cancer Res. 1997 Jun 1;57(11):2116-20. PMID: 9187106.
18. Björkqvist AM, Tammilehto L, Nordling S, Nurminen M, Anttila S, Mattson K, et al. Comparison of DNA copy number changes in malignant mesothelioma, adenocarcinoma and large-cell anaplastic carcinoma of the lung. Br J Cancer. 1998;77(2):260-9. doi: 10.1038/bjc.1998.42. PMID: 9460997; PMCID: PMC2151225.
19. Garnis C, Lockwood WW, Vucic E, Ge Y, Girard L, Minna JD, et al. High resolution analysis of non-small cell lung cancer cell lines by whole genome tiling path array CGH. Int J Cancer. 2006 Mar 15;118(6):1556-64. doi: 10.1002/ijc.21491. PMID: 16187286.
20. Luk C, Tsao MS, Bayani J, Shepherd F, Squire JA. Molecular cytogenetic analysis of non-small cell lung carcinoma by spectral karyotyping and comparative genomic hybridization. Cancer Genet Cytogenet. 2001 Mar;125(2):87-99. doi: 10.1016/s0165-4608(00)00363-0. PMID: 11369051.
21. Petersen I, Bujard M, Petersen S, Wolf G, Goeze A, Schwendel A, et al. Patterns of chromosomal imbalances in adenocarcinoma and squamous cell carcinoma of the lung. Cancer Res. 1997 Jun 15;57(12): 2331-5. PMID: 9192802.
22. Testa JR, Siegfried JM, Liu Z, Hunt JD, Feder MM, Litwin S, et al. Cytogenetic analysis of 63 non-small cell lung carcinomas: recurrent chromosome alterations amid frequent and widespread genomic upheaval. Genes Chromosomes Cancer. 1994 Nov;11(3):178-94. doi: 10.1002/gcc.2870110307. PMID: 7530487.
23. Weir BA, Woo MS, Getz G, Perner S, Ding L, Beroukhim R, et al. Characterizing the cancer genome in lung adenocarcinoma. Nature. 2007 Dec 6;450(7171):893-8. doi: 10.1038/nature06358. PMID: 17982442; PMCID: PMC2538683.
24. Tonon G, Wong KK, Maulik G, Brennan C, Feng B, Zhang Y, et al. High-resolution genomic profiles of human lung cancer. Proc Natl Acad Sci U S A. 2005 Jul 5;102(27):9625-30. doi: 10.1073/pnas.0504126102. PMID: 15983384; PMCID: PMC1160520.
25. Zhao X, Weir BA, LaFramboise T, Lin M, Beroukhim R, Garraway L, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. Cancer Res. 2005 Jul 1;65(13):5561-70. doi: 10.1158/0008-5472.CAN-04-4603. PMID: 15994928.
26. Giles Doran C, Pennington SR. Copy number alteration signatures as biomarkers in cancer: a review. Biomark Med. 2022 Apr;16(5):371-86. doi: 10.2217/bmm-2021-0476. PMID: 35195030.
27. Foster JM, Oumie A, Togneri FS, Vasques FR, Hau D, Taylor M, et al. Cross-laboratory validation of the OncoScan® FFPE Assay, a multiplex tool for whole genome tumour profiling. BMC Med Genomics. 2015 Feb 18;8:5. doi: 10.1186/s12920-015-0079-z. PMID: 25889064; PMCID: PMC4342810.
28. Hauenstein J, Liebenberg AP, Matthews B, O'Hare C, Thompson KS, Phillips CN, et al. Comparison of Genomic Coverage Using Affymetrix OncoScan Array and Illumina TruSight Tumor 170 NGS Panel for Detection of Copy Number Abnormalities in Clinical GBM Specimens. Cancer Genet. 2017 Aug;214:33. doi:10.1016/j.cancergen.2017.04.005.
29. Cho JW, Cao J, Hemberg M. Joint analysis of mutational and transcriptional landscapes in human cancer reveals key perturbations during cancer evolution. Genome Biol. 2024 Mar 8;25(1):65. doi: 10.1186/s13059-024-03201-1. PMID: 38459554; PMCID: PMC10921788.
30. Cornish AJ, Gruber AJ, Kinnersley B, Chubb D, Frangou A, Caravagna G, et al. The genomic landscape of 2,023 colorectal cancers. Nature.

2024 Sep;633(8028):127-136. doi: 10.1038/s41586-024-07747-9. PMID: 39112709; PMCID: PMC11374690.
31. Wang Z, Luo M, Liang Q, Zhao K, Hu Y, Wang W, et al. Landscape of enhancer disruption and functional screen in melanoma cells. Genome Biol. 2023 Oct 30;24(1):248. doi: 10.1186/s13059-023-03087-5. PMID: 37904237; PMCID: PMC10614365.
32. Krijgsman O, Carvalho B, Meijer GA, Steenbergen RD, Ylstra B. Focal chromosomal copy number aberrations in cancer-Needles in a genome haystack. Biochim Biophys Acta. 2014 Nov;1843(11): 2698-2704. doi: 10.1016/j.bbamcr.2014.08.001. PMID: 25110350.
33. Leary RJ, Lin JC, Cummins J, Boca S, Wood LD, Parsons DW, et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci U S A. 2008 Oct 21;105(42):16224-9. doi: 10.1073/pnas.0808041105. PMID: 18852474; PMCID: PMC2571022.
34. Deheuninck J, Luo K. Ski and SnoN, potent negative regulators of TGF-beta signaling. Cell Res. 2009 Jan;19(1):47-57. doi: 10.1038/cr.2008.324. PMID: 19114989; PMCID: PMC3103856.
35. Ma F, Ding MG, Lei YY, Luo LH, Jiang S, Feng YH, et al. SKIL facilitates tumorigenesis and immune escape of NSCLC via upregulating TAZ/autophagy axis. Cell Death Dis. 2020 Dec 2;11(12):1028. doi: 10.1038/s41419-020-03200-7. PMID: 33268765; PMCID: PMC7710697.
36. Pullamsetti SS, Banat GA, Schmall A, Szibor M, Pomagruk D, Hänze J, et al. Phosphodiesterase-4 promotes proliferation and angiogenesis of lung cancer by crosstalk with HIF. Oncogene. 2013 Feb 28;32(9): 1121-34. doi: 10.1038/onc.2012.136. PMID: 22525277.
37. Jensen MR, Stoltze U, Hansen TVO, Bak M, Sehested A, Rechnitzer C, et al. 9p21.3 Microdeletion involving CDKN2A/2B in a young patient with multiple primary cancers and review of the literature. Cold Spring Harb Mol Case Stud. 2022 Jun 22;8(4):a006164. doi: 10.1101/mcs.a006164. PMID: 35422439; PMCID: PMC9235845.
38. Jiang J, Gu Y, Liu J, Wu R, Fu L, Zhao J, et al. Coexistence of p16/CDKN2A homozygous deletions and activating EGFR mutations in lung adenocarcinoma patients signifies a poor response to EGFR-TKIs. Lung Cancer. 2016 Dec;102:101-107. doi: 10.1016/j.lungcan.2016.10.015. PMID: 27987577.
39. Jin JO, Puranik N, Bui QT, Yadav D, Lee PC. The ubiquitin system: an emerging therapeutic target for lung cancer. Int J Mol Sci. 2021 Sep 6;22(17):9629. doi: 10.3390/ijms22179629. PMID: 34502538; PMCID: PMC8431782.
40. Hu J, Yang D, Zhang H, Liu W, Zhao Y, Lu H, et al. USP22 promotes tumor progression and induces epithelial-mesenchymal transition in lung adenocarcinoma. Lung Cancer. 2015 Jun;88(3):239-45. doi: 10.1016/j.lungcan.2015.02.019. PMID: 25907317.
41. McFarlane C, McFarlane S, Paul I, Arthur K, Scheaff M, Kerr K, et al. The deubiquitinating enzyme USP17 is associated with non-small cell lung cancer (NSCLC) recurrence and metastasis. Oncotarget. 2013 Oct;4(10):1836-43. doi: 10.18632/oncotarget.1282. PMID: 24123619; PMCID: PMC3858568.
42. Pan J, Deng Q, Jiang C, Wang X, Niu T, Li H, et al. USP37 directly deubiquitinates and stabilizes c-Myc in lung cancer. Oncogene. 2015 Jul 23;34(30):3957-67. doi: 10.1038/onc.2014.327. PMID: 25284584.
43. Sun XX, He X, Yin L, Komada M, Sears RC, Dai MS. The nucleolar ubiquitin-specific protease USP36 deubiquitinates and stabilizes c-Myc. Proc Natl Acad Sci U S A. 2015 Mar 24;112(12):3734-9. doi: 10.1073/pnas.1411713112. PMID: 25775507; PMCID: PMC4378440.
44. Zhang C, Lu J, Zhang QW, Zhao W, Guo JH, Liu SL, et al. USP7 promotes cell proliferation through the stabilization of Ki-67 protein in non-small cell lung cancer cells. Int J Biochem Cell Biol. 2016 Oct;79:209-21. doi: 10.1016/j.biocel.2016.08.025. PMID: 27590858.
45. Kane RC, Bross PF, Farrell AT, Pazdur R. Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. Oncologist. 2003;8(6):508-13. doi: 10.1634/theoncologist.8-6-508. PMID: 14657528.
46. Drakos E, Singh RR, Rassidakis GZ, Schlette E, Li J, Claret FX, et al. Activation of the p53 pathway by the MDM2 inhibitor nutlin-3a overcomes BCL2 overexpression in a preclinical model of diffuse large B-cell lymphoma associated with t(14;18)(q32;q21). Leukemia. 2011 May;25(5):856-67. doi: 10.1038/leu.2011.28. PMID: 21394100; PMCID: PMC3094765.
47. Hai J, Sakashita S, Allo G, Ludkovski O, Ng C, Shepherd FA, et al. Inhibiting MDM2-p53 interaction suppresses tumor growth in patient-derived non-small cell lung cancer xenograft models. J Thorac Oncol. 2015 Aug;10(8):1172-80. doi: 10.1097/JTO.0000000000000584. PMID: 26200271.
48. Vu B, Wovkulich P, Pizzolato G, Lovey A, Ding Q, Jiang N, et al. Discovery of RG7112: a small-molecule MDM2 inhibitor in clinical development. ACS Med Chem Lett. 2013 Apr 2;4(5):466-9. doi: 10.1021/ml4000657. PMID: 24900694; PMCID: PMC4027145.
49. Chang YS, Graves B, Guerlavais V, Tovar C, Packman K, To KH, et al. Stapled α-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. Proc Natl Acad Sci U S A. 2013 Sep 3;110(36):E3445-54. doi: 10.1073/pnas.1303002110. PMID: 23946421; PMCID: PMC3767549.
50. El Osta B, Behera M, Kim S, Berry LD, Sica G, Pillai RN, et al. Characteristics and outcomes of patients with metastatic KRAS-Mutant lung adenocarcinomas: the Lung Cancer Mutation Consortium experience. J Thorac Oncol. 2019 May;14(5):876-889. doi: 10.1016/j.jtho.2019.01.020. PMID: 30735816; PMCID: PMC8108452.
51. Wagner PL, Stiedl AC, Wilbertz T, Petersen K, Scheble V, Menon R, et al. Frequency and clinicopathologic correlates of KRAS amplification in non-small cell lung carcinoma. Lung Cancer. 2011 Oct;74(1):118-23. doi: 10.1016/j.lungcan.2011.01.029. PMID: 21477882.
52. Sasaki H, Hikosaka Y, Kawano O, Moriyama S, Yano M, Fujii Y. Evaluation of Kras gene mutation and copy number gain in non-small cell lung cancer. J Thorac Oncol. 2011 Jan;6(1):15-20. doi: 10.1097/JTO.0b013e31820594f0. PMID: 21150464.
53. Jänne PA, Riely GJ, Gadgeel SM, Heist RS, Ou SI, Pacheco JM, et al. Adagrasib in non-small-cell lung cancer harboring a KRASG12C mutation. N Engl J Med. 2022 Jul 14;387(2):120-131. doi: 10.1056/NEJMoa2204619. PMID: 35658005.
54. Skoulidis F, Li BT, Dy GK, Price TJ, Falchook GS, Wolf J, et al. Sotorasib for lung cancers with KRAS p.G12C mutation. N Engl J Med. 2021 Jun 24;384(25):2371-81. doi: 10.1056/NEJMoa2103695. PMID: 34096690; PMCID: PMC9116274.
55. Blanco R, Iwakawa R, Tang M, Kohno T, Angulo B, Pio R, et al. A gene-alteration profile of human lung cancer cell lines. Hum Mutat. 2009 Aug;30(8):1199-206. doi: 10.1002/humu.21028. PMID: 19472407; PMCID: PMC2900846.
56. Garnis C, Lockwood WW, Vucic E, Ge Y, Girard L, Minna J, et al. High resolution analysis of non-small cell lung cancer cell lines by whole genome tiling path array CGH. Int J Cancer. 2006 Mar 15;118(6):1556-64. doi: 10.1002/ijc.21491. PMID: 16187286.

SUPPLEMENTARY TABLES

Table S1. Genomic copy number (CN) alterations in GL01 and A549 cells identified using OncoScan arrays, arranged by Chromosome (Chr)

Cells	CN	Type	Chr	Cytoband	Start	End	Size (kbp)	Marker Count	Gene Count
A549	3.00	Gain	1	p36.32-36.33	818,812	3,719,224	2,900	288	101
GL01	3.00	Gain	1	q21.1-44	143,549,778	248,918,679	105,369	8,079	1,348
A549	3.00	Gain	1	q21.1-44	143,549,778	248,918,679	105,369	8,079	1,348
A549	3.00	Gain	2	p22.3-25.3	21,494	35,705,811	35,684	2,841	278
GL01	3.00	Gain	2	p22.3-25.3	21,494	35,793,078	35,772	2,844	278
A549	4.00	Gain	2	p16.3-22.3	35,745,263	51,407,793	15,663	1,252	108
GL01	4.00	Gain	2	p16.322.3	35,796,918	51,065,194	15,268	1,226	108
GL01	3.00	Gain	2	p14-16.3	51,071,789	66,223,903	15,152	1,145	87
A549	3.00	Gain	2	p14-16.3	51,417,597	66,130,772	14,713	1,119	87
A549	3.00	Gain	2	p11.2-12	77,419,326	88,860,933	11,442	712	88
GL01	3.00	Gain	2	p11.2-12	77,436,092	88,825,618	11,390	707	88
GL01	9.00	Gain	2	p11.2	88,839,118	89,127,208	288	41	0
A549	7.00	Gain	2	p11.2	88,864,261	89,127,208	263	37	0
GL01	3.00	Gain	2	p11.2-q11.1	89,127,578	95,497,204	6,370	72	22
GL01	3.00	Gain	3	p25.3-26.3	21,733	9,811,788	9,790	739	53
A549	3.00	Gain	3	p25.3-26.3	21,733	9,811,788	9,790	739	53
GL01	3.00	Gain	3	p22.1-25.3	10,170,748	39,894,328	29,724	2,248	202
A549	3.00	Gain	3	p22.1-25.3	10,180,701	39,894,328	29,714	2,238	202
A549	1.00	Loss	3	p14.2	60,087,427	61,253,897	1,166	152	2
GL01	1.00	Loss	3	p14.2	60,455,464	61,173,360	718	100	2
GL01	3.00	Gain	3	p14.1-q13.31	69,482,713	116,562,561	47,080	3,430	203
A549	3.00	Gain	3	p14.1-q13.31	69,482,713	116,562,561	47,080	3,430	203
A549	1.00	Loss	3	q13.31	116,570,247	116,976,364	406	26	3
GL01	1.00	Loss	3	q13.31	116,570,247	117,070,604	500	31	3
A549	3.00	Gain	3	q13.31-21.2	117,010,382	124,952,927	7,943	625	76
GL01	3.00	Gain	3	q13.31-21.2	117,082,201	124,952,927	7,871	620	76
A549	3.00	Gain	3	q23-26.1	141,261,077	166,102,852	24,842	1,729	161
GL01	3.00	Gain	3	q23-26.1	141,278,783	166,102,852	24,824	1,728	161
GL01	3.00	Gain	3	q26.2	169,000,309	170,396,290	1,396	139	15
A549	3.00	Gain	3	q26.2	169,000,309	170,590,642	1,590	152	19
A549	3.00	Gain	3	q26.31	172,713,229	173,226,893	514	35	2
A549	1.00	Loss	4	p16.3	1,930,877	2,292,851	362	32	11
A549	3.00	Gain	5	p15.33	38,141	1,391,254	1,353	128	32
GL01	4.00	Gain	5	p11-15.33	38,141	46,401,169	46,363	3,235	276
A549	4.00	Gain	5	p11-15.33	1,404,433	46,401,169	44,997	3,107	244
GL01	2.50	Gain	5	q11.2-12.1	58,416,723	59,643,060	1,226	112	6
A549	1.50	Loss	5	q12.1	62,120,841	62,602,489	482	32	5
A549	3.00	Gain	5	q32	150,164,482	150,358,353	194	20	5
A549	3.00	Gain	5	q35.2-35.3	174,355,328	179,334,357	4,979	365	77
GL01	2.50	Gain	5	q35.2	174,363,979	174,624,114	260	18	1
GL01	3.00	Gain	5	q35.2-35.3	174,638,392	181,271,311	6,633	487	123
A549	2.50	Gain	5	q35.3	179,349,821	179,647,385	298	21	5
A549	3.00	Gain	5	q35.3	179,665,140	181,271,311	1,606	120	41
GL01	1.00	Loss	6	p21.33	31,211,764	31,331,648	120	18	2

Table S1. Genomic copy number (CN) alterations in GL01 and A549 cells identified using OncoScan arrays, arranged by Chromosome (Chr) *(continued)*

Cells	CN	Type	Chr	Cytoband	Start	End	Size (kbp)	Marker Count	Gene Count
A549	1.00	Loss	6	p21.31-21.32	33,493,027	33,627,622	135	26	4
A549	2.50	Gain	6	p21.31	33,629,302	34,194,968	566	44	11
A549	1.50	Loss	6	q26	160,959,810	161,931,626	972	167	6
A549	1.00	Loss	6	q26	161,944,869	161,985,044	40	18	1
A549	1.00	Loss	6	q27	169,189,830	170,138,127	948	55	10
A549	3.00	Gain	7	p22.3-q22.1	41,421	100,419,831	100,378	7,237	805
GL01	3.00	Gain	7	p22.3-q33	41,421	133,869,745	133,828	9,911	1,092
A549	2.50	Gain	7	q22.1	100,435,042	101,147,131	712	66	34
A549	3.00	Gain	7	q22.1	102,461,717	134,141,648	31,680	2,529	229
GL01	5.00	Gain	7	q33	133,895,694	134,677,696	782	55	7
A549	5.00	Gain	7	q33	134,163,505	134,677,696	514	35	6
GL01	3.00	Gain	7	q33-36.3	134,697,231	159,325,753	24,629	2,005	250
A549	3.00	Gain	7	q33-36.3	134,697,231	159,325,753	24,629	2,005	250
A549	3.00	Gain	8	p21.3-23.3	222,417	22,250,834	22,028	2,169	211
GL01	3.00	Gain	8	p11.22-23.3	222,417	39,257,430	39,035	3,492	356
A549	3.00	Gain	8	p11.22-21.3	22,474,361	39,366,137	16,892	1,307	143
GL01	0.00	Loss	8	p11.22	39,281,679	39,539,496	258	18	3
A549	0.00	Loss	8	p11.22	39,382,315	39,648,307	266	19	4
GL01	3.00	Gain	8	p11.22-q21.3	39,554,938	89,305,466	49,751	3,544	261
A549	3.00	Gain	8	p11.22-q11.21	39,665,094	47,343,547	7,678	435	43
A549	2.50	Gain	8	q11.21	47,368,914	47,860,871	492	68	3
A549	3.00	Gain	8	q11.21-21.3	47,863,811	89,254,132	41,390	3,030	216
GL01	3.00	Gain	8	q21.3-24.13	89,882,782	124,922,451	35,040	3,252	212
A549	3.00	Gain	8	q21.3-24.13	89,933,828	124,925,518	34,992	3,247	211
GL01	4.00	Gain	8	q24.13-24.3	124,925,518	145,067,348	20,142	1,604	218
A549	4.00	Gain	8	q24.13-24.3	124,947,521	139,638,561	14,691	1,154	73
A549	3.00	Gain	8	q24.3	139,641,860	139,897,240	255	18	2
A549	4.00	Gain	8	q24.3	139,911,832	141,395,440	1,484	146	14
A549	3.00	Gain	8	q24.3	141,419,373	141,766,385	347	22	3
A549	4.00	Gain	8	q24.3	141,782,700	143,037,818	1,255	89	24
A549	2.50	Gain	8	q24.3	143,044,827	143,507,042	462	30	15
A549	3.00	Gain	8	q24.3	143,527,387	144,231,771	704	77	42
A549	4.00	Gain	8	q24.3	144,280,183	144,735,875	456	46	34
GL01	3.00	Gain	9	p21.3-24.3	204,738	21,828,111	21,623	1,858	110
A549	3.00	Gain	9	p21.3-24.3	204,738	21,828,111	21,623	1,858	110
GL01	0.00	Loss	9	p21.3	21,833,660	22,119,196	286	160	5
A549	0.00	Loss	9	p21.3	21,833,660	22,119,196	286	160	5
A549	3.00	Gain	9	p21.3-q21.32	22,140,225	83,691,298	61,551	2,277	293
GL01	3.00	Gain	9	p21.3-q21.32	22,140,225	83,847,074	61,707	2,287	296
A549	3.00	Gain	10	p11.21-15.3	80,130	34,906,478	34,826	2,593	253
GL01	3.00	Gain	10	p15.3-q26.3	80,130	133,620,799	133,541	9,695	1,104
A549	2.50	Gain	10	p11.21	34,916,714	35,494,128	577	39	4
A549	3.00	Gain	10	p11.21-q22.2	35,503,645	73,627,791	38,124	2,465	277
A549	2.50	Gain	10	q22.2	73,667,019	74,683,339	1,016	70	17
A549	3.00	Gain	10	q22.2-24.31	74,709,093	100,700,729	25,992	2,160	254
A549	2.50	Gain	10	q24.31	100,723,374	100,984,457	261	19	5

Table S1. Genomic copy number (CN) alterations in GL01 and A549 cells identified using OncoScan arrays, arranged by Chromosome (Chr) *(continued)*

Cells	CN	Type	Chr	Cytoband	Start	End	Size (kbp)	Marker Count	Gene Count
A549	3.00	Gain	10	q24.31-26.3	100,996,022	133,620,799	32,625	2,349	293
GL01	3.00	Gain	11	p15.5-q13.3	192,764	69,633,807	69,441	4,787	1,061
A549	3.00	Gain	11	p15.5-q14.3	192,764	92,476,687	92,284	6,539	1,260
GL01	4.00	Gain	11	q13.3	69,634,765	69,665,423	31	35	1
GL01	3.00	Gain	11	q13.3-14.3	69,665,816	92,599,875	22,934	1,725	198
GL01	3.00	Gain	12	p13.33-q24.33	80,234	133,241,529	133,161	10,345	1,391
A549	3.00	Gain	12	p13.33-q24.33	80,234	133,241,529	133,161	10,345	1,391
A549	3.00	Gain	14	q11.2-12	19,750,924	24,249,309	4,498	440	150
GL01	3.00	Gain	14	q11.2-32.33	19,750,924	105,767,230	86,016	6,562	917
A549	2.50	Gain	14	q12	24,266,418	24,691,739	425	30	18
A549	3.00	Gain	14	q12-32.11	24,706,755	90,370,021	65,663	4,887	452
A549	3.00	Gain	14	q32.11-32.33	90,662,551	105,760,702	15,098	1,186	292
A549	4.00	Gain	14	q32.33	105,767,230	106,388,139	621	105	7
GL01	5.00	Gain	14	q32.33	105,775,617	106,383,907	608	103	7
GL01	3.00	Gain	14	q32.33	106,388,139	106,873,815	486	79	2
A549	3.00	Gain	14	q32.33	106,398,004	106,873,815	476	78	2
GL01	4.00	Gain	15	q11.1-11.2	19,956,119	24,747,241	4,791	134	59
A549	4.00	Gain	15	q11.1-11.2	19,956,119	24,747,241	4,791	134	59
A549	4.00	Gain	15	q12-14	26,475,159	35,364,641	8,889	556	87
GL01	4.00	Gain	15	q12-14	26,475,159	35,424,734	8,950	560	89
GL01	4.00	Gain	15	q21.2-21.3	49,810,482	53,730,468	3,920	312	40
A549	4.00	Gain	15	q21.2-21.3	49,810,482	53,730,468	3,920	312	40
GL01	4.00	Gain	15	q21.3-24.1	55,032,190	73,665,458	18,633	1,406	202
A549	4.00	Gain	15	q21.3-24.1	55,032,190	73,665,458	18,633	1,406	202
GL01	4.00	Gain	15	q24.1-26.3	74,763,008	101,857,114	27,094	2,183	307
A549	4.00	Gain	15	q24.1-26.3	74,763,008	101,857,114	27,094	2,183	307
A549	3.00	Gain	16	p12.1-13.3	33,887	27,624,519	27,591	2,110	462
GL01	3.00	Gain	16	p13.3-q24.3	33,887	90,091,597	90,058	5,834	1,141
A549	4.00	Gain	16	p11.2-12.1	27,636,509	28,616,904	980	78	16
A549	2.50	Gain	16	p11.2	28,617,824	29,303,678	686	45	19
A549	3.00	Gain	16	p11.2-q24.3	29,308,601	90,091,597	60,783	3,601	647
A549	3.00	Gain	17	p13.3	497,719	1,307,207	809	54	13
GL01	4.00	Gain	17	p13.3-q25.3	497,719	82,305,551	81,808	6,413	1,552
A549	4.00	Gain	17	p13.2-13.3	1,323,447	5,179,933	3,856	325	118
A549	3.00	Gain	17	p11.1-13.2	5,182,562	22,718,556	17,536	1,362	329
A549	4.00	Gain	17	q11.1-12	26,999,915	38,442,381	11,442	848	206
A549	3.00	Gain	17	q12	38,455,815	38,837,804	382	72	15
A549	4.00	Gain	17	q12	38,846,880	39,677,589	831	115	26
A549	2.50	Gain	17	q12	39,678,808	39,757,620	79	79	5
A549	4.00	Gain	17	q12-25.3	39,773,563	82,305,551	42,532	3,558	843
A549	1.00	Loss	18	q23	79,007,070	79,313,322	306	21	1
A549	2.50	Gain	18	q23	79,326,759	80,249,901	923	75	14
GL01	4.00	Gain	19	p11-12	21,831,161	24,361,518	2,530	164	30
A549	4.00	Gain	19	p11-12	21,831,161	24,361,518	2,530	164	30
A549	5.00	Gain	19	q11-13.12	27,263,665	35,626,525	8,363	661	95
GL01	5.00	Gain	19	q11-13.2	27,263,665	42,800,407	15,537	1,421	335

Table S1. Genomic copy number (CN) alterations in GL01 and A549 cells identified using OncoScan arrays, arranged by Chromosome (Chr) *(continued)*

Cells	CN	Type	Chr	Cytoband	Start	End	Size (kbp)	Marker Count	Gene Count
A549	4.00	Gain	19	q13.12-13.43	35,633,274	58,581,872	22,949	2,052	965
GL01	4.00	Gain	19	q13.2-13.43	42,817,913	58,581,872	15,764	1,292	725
A549	4.00	Gain	20	p12.1-13	88,453	14,806,114	14,718	1,396	164
GL01	4.00	Gain	20	p12.1-13	88,453	14,812,769	14,724	1,397	164
GL01	1.00	Loss	20	p12.1	14,921,130	15,052,199	131	53	2
A549	1.00	Loss	20	p12.1	14,921,130	15,057,252	136	54	2
A549	3.00	Gain	20	p12.1	15,058,274	15,321,095	263	108	1
GL01	3.00	Gain	20	p12.1	15,167,738	15,321,095	153	64	1
GL01	4.00	Gain	20	p11.1-12.1	15,324,695	26,328,619	11,004	996	100
A549	4.00	Gain	20	p11.1-12.1	15,324,695	26,328,619	11,004	996	100
A549	2.50	Gain	20	q11.1-11.21	30,284,480	32,158,445	1,874	100	33
GL01	3.00	Gain	20	q11.1-13.33	30,284,480	64,281,110	33,997	2,493	514
A549	3.00	Gain	20	q11.21-13.33	32,167,895	64,281,110	32,113	2,393	481
GL01	3.00	Gain	21	p11.2-q21.1	8,759,482	17,886,994	9,128	286	38
A549	3.00	Gain	21	p11.2-q21.1	8,759,482	17,886,994	9,128	286	38
GL01	1.00	Loss	21	q21.1	17,906,297	18,204,879	299	21	1
A549	1.00	Loss	21	q21.1	17,906,297	18,204,879	299	21	1
A549	3.00	Gain	21	q21.1-22.11	18,222,841	34,310,923	16,088	1,108	132
GL01	3.00	Gain	21	q21.1-22.3	18,222,841	46,677,698	28,455	2,040	330
A549	2.50	Gain	21	q22.11-22.13	34,326,877	36,719,738	2,393	175	27
A549	3.00	Gain	21	q22.13-22.2	36,734,011	40,840,312	4,106	340	39
A549	2.50	Gain	21	q22.2-22.3	40,856,590	46,677,698	5,821	417	133
GL01	3.00	Gain	X	p22.33	261,275	2,751,360	2,490	107	22
A549	3.00	Gain	X	p22.33	261,275	2,789,693	2,528	110	23
A549	2.00	Gain	X	p22.33-q28	2,792,799	155,699,751	152,907	11,058	1,133
GL01	1.00	Gain	Y	p11.2-q12	2,792,122	26,653,788	23,862	613	116

Table S2. List of genes involved in copy number aberrations (CNAs) of GL01 and A549 that are known to be cancer drivers, oncogenes, or tumor suppressors in lung adenocarcinoma, identified using CancerMine, a literature-mined database of drivers, oncogenes, and tumor suppressors in various cancers[16]

Genes found in CNAs of both GL01 and A549 (total: 317 genes)					
EGFR	ASPM	HSD17B6	METTL1	CPSF4	VRK1
KRAS	CBLC	DEPTOR	MDM2	CPM	USP22
ALK	TGIF2	SOCS2	MAP2K5	SLC13A5	UPP1
TP53	TIAM1	KIF14	MAP2K1	STRIP2	UBE2V1
RET	SFTPC	MAGI2	KTN1	CHCHD2	U2AF1
BRAF	RRM2	CFTR	KRT16	SMYD2	TWIST1
MYC	BACH1	RHOBTB2	ITGB1	SLC39A5	TPM2
NKX2-1	NRG1	PDZD2	IRF9	CLDN3	TPD52L2
PTGDR	KIF5B	DHRS2	ANXA8	TTC9	HNF1A
ERBB2	GATA3	CEBPA	IFITM3	FLVCR2	SOX12
CD274	AKT1	CDK5	HSF2BP	GREM1	SOS1
EML4	CDH1	ACVRL1	HRAS	CYCS	SNAI1
MET	ZEB1	SOX6	HOXC13	RDM1	SNAI2
CDKN2A	PTPRO	IL33	HOXB3	TC2N	PRMT5
METTL3	PRKAR1A	NDRG2	HNRNPA1	TP53I3	SHOX2
MAP2K4	AXL	IRX1	FOXA1	CMTM6	SRSF1
MYBL2	PGM5	BRD7	HMGN1	S100A14	SFRP4
FAM83A	PDK4	TRIM55	IQANK1	CCDC6	SEMA4B
RIT1	PDCD4	HDAC9	GPX2	PDCD1LG2	SAA1
FGFR2	NR5A2	ZFP36	GPR87	EIF4A3	ROBO1
NF1	NFKBIA	VPS33B	AMACR	LRRK2	RASGRF1
FMO2	ARRB1	UBQLN1	GPR37	CENPF	RAF1
SIRT3	MCPH1	TNFRSF6B	ALOX5	SYT12	PTPRD
PTDSS1	SMAD3	TMPRSS2	ALCAM	UBE2S	PIWIL1
NTRK1	IRAG2	TBX4	AKT2	MYG1	ATR
HOXA11	JAK2	TAT	AKR1B10	SUV39H2	KAT2B
ELF3	NOD2	STAT3	FOS	CD8A	ASPH
AGR2	HOXA4	BTG2	FASN	ECD	KMT2D
TRIB2	GNG11	SPN	EZH2	TMED2	EIF6
IGF2BP3	FZD3	SPINT2	ERG	MELK	IGF1R
AHNAK2	FOXP1	SFTPA1	EPHA3	TRIM7	HSP90AA1
CCND1	FBLN5	SCAMP3	MARK2	CCT3	GNAS
RBL2	FABP4	RELB	EIF4EBP1	ZSWIM1	FOXM1
PTEN	EMX2	BCL2L1	EIF3H	PHF20	FGFR1
MEN1	DUSP6	PTN	EEF1A2	TSG101	EPHB4
DOK2	SNX20	PTK2	CISD1	CCNB2	EFNA3
ZNF671	RASL12	PRC1	TARS2	CBX4	E2F1
PAX6	PCNP	CTSA	RALGPS2	GOLPH3	ATAD2
P4HA1	NLRC3	B4GALT1	C1orf43	RAI14	SGO1
NUCB2	DNMT3B	PLK1	IGF2BP1	F11R	DDIT4
MYCN	MTUS1	PLAGL2	HILPDA	PARVA	OTUB1
MYBPH	TET1	PKP2	C1orf74	NETO2	NCAPG2
GSDMC	ATMIN	PI3	CYP24A1	CDCA5	AATF
MCM4	DLC1	SERPINA1	DDX59	CDCA4	CEACAM1
KITLG	DKK1	AVP	UBE2T	MRPL42	LMTK2
HMGA2	ENKUR	PFKFB3	GPRIN1	IRX5	CD44
GTF2E2	CUEDC2	PAWR	DCBLD2	OPN3	GGA2
ERBB3	SPNS2	P2RY1	GINS2	ERVW-1	FOSL1
KIF18B	ADPRH	NUMBL	C1GALT1	SPAG5	DEGS1
FANCI	CPED1	NEK2	FAM111B	RBMS3	TSPAN4
VIRMA	AMBRA1	NEDD4	FAM72A	ARHGEF5	RUNX1
TNS4	CTSF	MYEOV	E2F7	ZNF143	RPS27
HKDC1	BTNL9	MUC1	CRK	XPO1	
Genes found in CNAs of A549 only (total: 31 genes)					
RBM10	TAFAZZIN	PIR	MID1	HAUS3	BEX1
ACSL4	MAOA	CYBB	KLF8	MCTS1	USP51
CDK16	ARHGAP6	BEX4	FOXP3	CAMK2A	NSD2
ARAF	GPC3	USP11	G6PD	ACE2	TNFRSF4
ECT2	GJB1	SMARCA1	RER1	THBS2	TIMP1
S100G					