Survival-Critical Genes Associated with Copy Number Alterations in Lung Adenocarcinoma

Chinthalapally V. Rao, Chao Xu, Mudassir Farooqui, Yuting Zhang, Adam S. Asch, and Hiroshi Y. Yamada

Center for Cancer Prevention and Drug Development, Department of Medicine, Hematology/Oncology Section, University of Oklahoma Health Sciences Center (OUHSC), Oklahoma City, OK 73104, USA; yuting-zhang@ouhsc.edu

Stephenson Cancer Center, University of Oklahoma Health Sciences Center (OUHSC), Oklahoma City, OK 73104, USA; adam-asch@ouhsc.edu

VA Medical Center, Oklahoma City, OK 73104, USA

Hudson College of Public Health, University of Oklahoma Health Sciences Center (OUHSC), Oklahoma City, OK 73104, USA; chao-xu@ouhsc.edu

Department of Neurology, University of Iowa Hospitals and Clinics, Iowa City, IA 52242, USA; mudassir-farooqui@uiowa.edu

Correspondence: cv-rao@ouhsc.edu (C.V.R.); hiroshi-yamada@ouhsc.edu (H.Y.Y.)

Simple Summary: Genomic instability affects cancer evolution and impacts carcinogenesis, therapy response, recurrence/prognosis, thus overall clinical outcomes. To comprehensively identify genes and pathways affecting genomic instability, we employed a novel data-mining strategy (Gene Expression to Copy Number Alterations; “GE-CNA” approach) and identified 1578 genes whose expression associates with Copy Number Alterations in human lung adenocarcinoma. Among the 1578 genes, we identified 39 as survival-critical. They represent potential targets for therapy development.

Abstract: Chromosome Instability (CIN) in tumors affects carcinogenesis, drug resistance, and recurrence/prognosis. Thus, it has a high impact on outcomes in clinic. However, how CIN occurs in human tumors remains elusive. Although cells with CIN (i.e., pre/early cancer cells) are proposed to be removed by apoptosis and/or a surveillance mechanism, this surveillance mechanism is poorly understood. Here we employed a novel data-mining strategy (Gene Expression to Copy Number Alterations [CNA]; “GE-CNA”) to comprehensively identify 1578 genes that associate with CIN, indicated by genomic CNA as its surrogate marker, in human lung adenocarcinoma. We found that (a) amplification/insertion CNA is facilitated by over-expressions of DNA replication stressor and suppressed by a broad range of immune cells (T-, B-, NK-cells, leukocytes), and (b) deletion CNA is facilitated by over-expressions of mitotic regulator genes and suppressed predominantly by leukocytes guided by leukocyte extravasation signaling. Among the 39 CNA- and survival-associated genes, the purine metabolism (PPAT, PAICS), immune-regulating CD4-LCK-MEC2C and CCL14-CCR1 axes, and ALOX5 emerged as survival-critical pathways. These findings revealed a broad role of the immune system in suppressing CIN/CNA and cancer development in lung, and identified components representing potential targets for future chemotherapy, chemoprevention, and immunomodulation approaches for lung adenocarcinoma.

Keywords: genomic instability; lung tumor; tumor heterogeneity; tumor genomics

1. Introduction

Tumor genomics data have revealed a wide variety of tumor heterogeneity, the underlying driver of which is genomic instability. There are two major modes of genomic instability: Microsatellite Instability (MIN) and Chromosome Instability (CIN). Although...
not mutually exclusive, MIN is generally associated with functional alteration in proteins involved in DNA replication and repair, while CIN is mainly associated with mitotic dysregulation. In a solid tumor, genomic instability, especially large-scale chromosomal alterations caused by CIN, can be indicated with Copy Number Alterations (CNA) as its surrogate indicator [1]. Meta analyses indicated that high CIN is a marker of poor prognosis in colorectal, breast and lung cancers [2,3], consistent with the notion that genomic instability aids tumor evolution and fitness for survival under the challenges of cancer therapies [4]. Transgenic mouse-based studies demonstrated that elevated CIN can cause various cancers, including sporadic tumors in lung and liver [5–7]. Induced high CIN in mice increases the risk of lung cancer recurrence [8]. Thus, CIN is involved in carcinogenesis, drug response, and cancer recurrence, influencing major aspects of cancer prevention and therapy in the clinic setting [9–11]. The notion to target CIN and/or aneuploidy has long been proposed.

However, genes that govern the degree of CIN and CNA in tumors have been difficult to identify in a comprehensive manner. As genes/pathways to “target” remain unclear, the “Targeting CIN” concept has not been efficiently translated in clinic. Some of the “frequently mutated genes” in a given cancer can also play a role in genomic instability, as observed in colon cancer [11]. Yet this is not always the case, or is hard to demonstrate, in tumors with few penetrating mutations. Since CIN-associated mitotic regulator genes, such as the spindle checkpoint components (e.g., Mad1, Mad2, BubR1), are not always heavily mutated in tumors with high CIN [12,13], the understanding of how tumor CIN occurs remains elusive. By comparing gene expression profiles of high CNA tumors and low CNA tumors (a CNA-GE approach), correlations between high mitotic regulator expression and low immune signature were demonstrated in 12 types of cancers [14]. However, this general correlation is yet to be translated to targeted therapies. As aneuploid effects are proposed to be highly organ-specific [15], detailed analysis of an organ of interest is needed to identify specific markers and pathways for therapeutic translation.

Based on the tumor development profile in CIN mouse models [5–7], we hypothesized that CIN most prominently affects tumor development in lung and liver.

Using a novel in-silico analysis covering all reported human genes, we set out to comprehensively identify human genes whose expression shows correlation with CNA in human lung adenocarcinoma (the gene expression to CNA; “GE-CNA” strategy), anticipating that this approach would directly illuminate the network of genes through which lung cancer develops high CIN, or the mechanism by which the body antagonizes lung cells with genomic instability.

2. Materials and Methods

We downloaded the Lung Adenocarcinoma (TCGA, PanCancer Atlas, 2018) datasets, including the gene expression profile (available number of samples \( n = 510 \)), copy number alterations \( (n = 511) \), mutation from whole exome sequencing \( (n = 566) \), patients’ survival and clinical data \( (n = 566) \), from cBioportal (https://www.cbioportal.org/study/summary?id=luad_tcga_pan_can_atlas_2018 accessed 1 May 2020) [16,17]. The Z-scores of mRNA expression (batch normalized by RSEM [18] from Illumina HiSeq_RNASeqV2) were used. The putative copy-number alteration (CNA) was estimated from GISTIC 2.0 [19], with 0 indicating neutral/no change, positive value indicating gain/amplification, and negative value indicating deletion. We analyzed all CNAs, amplification CNAs, and deletion CNAs.

There were 20,531 genes and 510 subjects in the downloaded gene expression file. We excluded 340 genes that were completely missing in all subjects, while the included genes were complete in all subjects. For each gene, we sorted its expression in 510 subjects and picked the subjects with the top and bottom ten expression values to form a high expression group and a low expression group. We extracted the number of CNA of subjects in the high and low expression groups from the downloaded CNA file, from which missing values were excluded. We used Student’s \( t \)-test to examine the difference of CNA counts in two groups at significance level of 0.05. Multiple-testing was adjusted by \( q \)-value [20].
The significant genes were further divided into two groups: higher expression that resulted in more CNAs and higher expression that resulted in fewer CNAs. The gene set enrichment analyses were conducted by IPA (Ingenuity Pathway Analysis, QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) ( Germantown, MD, USA) [21] at significance level of 0.05 after Benjamini–Hochberg correction [22]. The presented pathway graphs were generated by IPA as well.

The association between the gene alteration and the patients’ overall survival was examined by the Cox Proportional-Hazards Model, with adjustment of the patients’ age and tumor stage. The adjusted covariates were selected by their univariate Cox regression p-value < 0.05. All available variables were considered, such as age, race, and tumor stage. We combined race groups with small sample size and used race variable with two levels: White and Other. We combined sub-level of tumor stage under each stage of stage 1 to 4. The tumor stage with four levels was used in the analysis. Patients having missing value were excluded. The Hazard Ratio (HR) and p-value of the gene were reported. We followed the definition of “altered” subject in cBioportal: altered = CNA (type of high-level amplification or homozygous deletion) + mutation (from whole exome sequencing). If a subject has any CNA and mutation in the gene, the subject is defined as altered. Otherwise, the subject is defined as unaltered. The difference of gene expression level in the altered and unaltered groups was tested by the Wilcoxon rank sum test. The significance level was 0.05. Survival curves and boxplots by altered/unaltered group are presented accordingly. We implemented all statistical analyses using R (v4.0.3) and R packages.

3. Results

The analysis was performed as outlined in Figure 1A,B, aiming at identifying all genes whose expression levels show correlation to CNA with 95% confidence. This “GE-CNA” data-mining strategy initially indicated two groups of genes: (i) genes whose high expression correlates with a higher number of CNA (Figure S1A; Table S1), implicating the function as a facilitator of CIN and CNA, and (ii) genes whose high expression correlates with a lower number of CNA (Supplementary Figure S1B; Supplementary Table S2), implicating the function as a suppressor of CIN and CNA. Pathway-IPA analysis on the 492 group (i) CNA facilitator genes mapped their functions as mitotic regulator genes, including the Kinetochore Metaphase Signaling Pathway (Supplementary Figure S2A), Mitotic Roles of Polo-Like Kinase, Role of CHK Proteins in Cell Cycle Checkpoint Control, Cell Cycle: G2/M DNA Damage Checkpoint Regulation, Role of BRCA1 in DNA Damage Response, Hereditary Breast Cancer Signaling, Estrogen-mediated S-phase Entry, Protein Ubiquitination Pathway, ATM Signaling, Cell Cycle Control of Chromosomal Replication, and Mismatch Repair in Eukaryotes (Supplementary Table S1). In contrast, the functions of the 1086 group (ii) CNA suppressor genes were concentrated on immune functions, including the Th1 and Th2 Activation Pathway (Figure S2B), Complement System, Leukocyte Extravasation Signaling, Crosstalk between Dendritic Cells and Natural Killer Cells, Agranulocyte Adhesion and Diapedesis, Granulocyte Adhesion and Diapedesis, Th2 Pathway, Th1 Pathway, CTLA4 Signaling in Cytotoxic T Lymphocytes, Role of Pattern Recognition, Receptors in Recognition of Bacteria and Viruses, CD28 Signaling in T Helper Cells, and iCOS-iCOSL Signaling in T Helper Cells (Table S2). These results are in good agreement with Davoli et al. (2017), who employed the CNA-GE strategy (reverse of GE-CNA) and reported mitotic and immune signatures as traits of high or low CNA tumors.
For each gene, we tested the difference of clinical staging and gender distribution in included and non-included (control) group. Fisher’s exact test was used for categorical variable staging having 4 levels stage 1 to 4, while Chi-squared test was used for binary variable gender. After multiple testing correction by \( q \)-value, there are no significant differences in regards to clinical staging and gender (minimum adjusted \( p \)-value 0.095 and 0.452 respectively). In addition, we tested the difference of age using \( t \)-test. There are no significant differences in age.
The above initial analysis covered all CNA, both amplification/insertion and deletion. Next, we questioned whether amplification/insertion CNA and deletion CNA are differentially affected by different sets of genes. 161 genes in Table S3 [(iii) amplification/insertion CNA facilitator genes] were concentrated on the Role of BRCA1 in DNA Damage Response (Figure 2A), Hereditary Breast Cancer Signaling, and Cell Cycle Control of Chromosomal Replication, and included BLM, PHF10, CDC4, and CDK6. Their main functions are involved in the DNA replication and repair pathways; dysregulation causes MIN and/or DNA replication stress, which can lead to CIN [23]. Thus, amplification/insertion CNA is suggested to be predominantly driven by MIN or CIN caused by DNA replication stress. Individual genes known or unknown to affect genomic stability were also identified. For example, among the top eight high-significance genes (q-value < 0.005), dysregulation of KPNA2 (Karyopherin α2/Importin α) (q = 0.0009) has long been known to cause mitotic defects, and KPNA2 is reported as overexpressed in various cancers with poor prognosis [24]. A previously unidentified function of KLHL7 (Kelch Like Family Member 7) and KLHL11, components of the BCR (BTB-CUL3-RBX1) E3 ubiquitin ligase complex, in lung adenocarcinoma genomic instability was also suggested (q = 0.0003, 0.04, respectively). GINS1 (q = 0.0029) is a component of the GINS DNA replication initiation complex involved in reactivation from quiescence [25]. FAM126A (q = 0.0021) is downregulated by catenin, suggested to be a part of the beta-catenin/Lef signaling pathway. PRR19 was recently identified as a partner of cyclin-like CNTD1, and is required for timely DSB repair and the formation of crossover-specific recombination complexes [26].

(A) Role of BRCA1 in DNA damage response

Figure 2. Cont.
Figure 2. Different enrichments of CNA facilitator genes for amplification/insertion CNA and for deletion CNA. (A) IPA for amplification/insertion CNA facilitator genes. Replication stress-inducing genes and pathways (e.g., Role of BRCA1 in DNA Damage Response) are enriched. (B) IPA for deletion CNA facilitator genes. Kinetochore Metaphase signaling pathway is shown. Purple highlighting indicates particular genes with significant GE-CNA correlations and/or a cluster of such genes in the IPA pathways.

187 genes in Supplementary Table S4 [(iv) deletion CNA facilitator genes] were enriched with known mitotic regulators Kinetochore Metaphase Signaling Pathway (Figure 2B), Protein Ubiquitination Pathway, Cyclins and Cell Cycle Regulation, and Cell Cycle: G2/M DNA Damage Checkpoint Regulation, indicating that overexpression of mitotic regulator genes leads to deletion CNA, which also confirms the role of mitotic mis-regulations in CNA. Notably, PSMD14 deubiquitinase ($q = 3.63646 \times 10^{-8}$), a subunit of 26S proteasome, was recently identified as upregulated in NSCLC [27]. Seven other 26S/20S proteasome subunits were isolated ($q < 0.05$; PSMD12, PSMB4, PSMA6, PSMB6, PSMC1, PSMD3, and PSMB3), illuminating the importance of the 26/20S proteasome integrity in lung adenocarcinoma genome maintenance. KIF18B (Kinesin Family Member 18B) ($q = 0.0105$) forms a complex with KIF2C (Kinesin Family Member 2C) ($q = 0.0313$) and constitutes the microtubule plus-end depolymerizing activity during mitosis [28]. ZWINT (ZW10 Interacting Kinetochore Protein) ($q = 0.0126$) is a kinetochore protein involved in mitotic checkpoint [29]. Other kinetochore proteins NUF2 ($q = 0.0130$) and SPC24 ($q = 0.0244$), as well as centromere proteins CENPI, CENPA, and CENPM, and mitotic regulatory kinases AURKA, TTK, and CDK1, are identified ($q < 0.05$). Also notable were transcription regula-
tors (DPY30, MESP1, POU4F1, HOXC8, FOXK2, HOXC10, DLX5, E2F6, and SP8 \( q < 0.05 \)). DPY30 (Dpy-30 Histone Methyltransferase Complex Regulatory Subunit) \( q = 0.0109 \) is a core subunit of the SET1/MLL family of H3K4 methyltransferases that directly controls cell cycle regulators [30].

420 genes in Supplementary Table S5 [(v) amplification/insertion CNA suppressor genes] included a broad range of immune system genes. The IPA pathways involved include the iCOS-iCOSL Signaling in T Helper Cells (Figure 3A), Crosstalk between Dendritic Cells and Natural Killer Cells, Th1 and Th2 Activation Pathway, Systemic Lupus Erythematosus In B Cell Signaling Pathway, LPS/IL-1 Mediated Inhibition of RXR Function, Leukocyte Extravasation Signaling, Th2 Pathway, Th1 Pathway, Phenylethylamine Degradation I, and Xenobiotic Metabolism CAR Signaling Pathway. Among the top 36 high-significance genes \( q < 0.005 \), GSTM5 (glutathione S transferase mu 5) \( q = 0.0003 \) suggests the role of oxidative stress defense and detoxification in reducing tumor CNA. A decrease in GSTM5 was also observed in the sporadic lung cancer-prone genomic instability mouse model Sgo1\(^{-/-}\) [7]. Also notable are immunomodulatory surface receptors (CD37, CD22, LILRB3, TLR3) and other immune modulators (DPEP2, FGR, IL18) \( q < 0.005 \).

Figure 3. Cont.
Leukocyte extravasation signaling

Figure 3. Different enrichments of CNA suppressor genes for amplification/insertion CNA and for deletion CNA (A). CNA suppressor genes for amplification/insertion CNA include a wide variety of immune pathways, e.g., iCOS-iCOSL signaling in T-helper cells. (B) CNA suppressor genes for deletion CNA showed specific enrichment (BH adjusted \( p = 0.0017 \)) on Leukocyte Extravasation Signaling. The enrichment of Leukocyte Extravasation Signaling suggests that tumor cells with deletion CNA are recognized and removed by leukocytes guided by the signaling.

In contrast, 407 genes in Supplementary Table S6 [(vi) deletion CNA suppressor genes] showed specific enrichment (BH adjusted \( p = 0.0017 \)) on Leukocyte Extravasation Signaling (Figure 3B) including ARHGAP6, ARHGAP9, CDH5, CXCR4, GNAI2, ITGA1, ITK, JAM2, MMP19, NCF1, PECAM1, PIK3R5, RAC2, SPN, and TEC [31]. These genes control the movement of leukocytes from blood vessels towards the site of tissue damage or infection, suggesting that leukocytes guided by Leukocytes Extravasation Signaling may specifically target cells with deletion-CNA, which may be generated by overexpression of "deletion CNA facilitator" genes, or mitotic errors (Supplementary Table S4). In particular, the functions of the top 30 high-significance genes (\( q < 0.005 \)) are concentrated on protection at the cell surface (MUC17, TSPAN3, TFF2, MUCL3) and membrane-based signal transduction (CHRM1, IL5RA, CDH5, TLR2, PTPRQ).

To identify potential targets for therapeutic drug or intervention, we surveyed the gene lists and tested whether the gene expression level affected patients’ overall survival, adjusted with covariates for lung adenocarcinoma. The second screening identified 39 genes for which expression alterations correlate with survival (Figure 4A; Supplementary Table S7). Among the 39 genes, 29 genes indicated a Hazard Ratio (HR) > 1, for which expression alterations increase risk, while 10 genes indicated HR < 1. Among CNA facilitators, PPAT and PAICS both indicated HR > 1 and their altered expressions were overexpression; both are involved in purine biosynthesis and metabolism (Figure 4B,C). Previous studies indicated that PPAT and PAICS were highly expressed in various cancers at advanced stages and involved in progression, which led to proposals for PPAT and PAICS as therapeutic inhibition targets [32]. Among CNA suppressors, the CD4-LCK-MEF2C
axis and CCL14-CCR1 axis emerged as survival-critical pathways (Figure 4D). KLHL2 (Kelch Like Family Member 2; a ubiquitin ligase) indicated HR = 0.1, and decreased KLHL2 expression correlated with better prognosis ($p = 0.0387$) (Figure 4E). PRX (Periaxin, a key myelination protein) indicated HR = 2.3, and overexpression correlated with poor prognosis ($p = 0.0293$) (Figure 4F). These findings suggest that KLHL2 and/or PRX may be a novel inhibition target for therapeutic purposes. ALOX5/5-LOX (Arachidonate 5-lipoxygenase) emerged as del-CNA suppressor (HR = 0.2, $p = 0.0241$; Figure 4G).

**From Table S3: Amp-CIN facilitator**

| gene  | table | hr    | p value |
|-------|-------|-------|---------|
| PRR19 | 3     | 2.005502 | 0.048102 |
| MAPRE1 | 3     | 6.494738 | 0.009651 |
| FBXL19 | 3     | 0.117526 | 0.034106 |
| MRPS17 | 3     | 2.155393 | 0.00229 |
| EXOSC10 | 3     | 2.615234 | 0.035744 |
| POUAF1 | 3     | 2.743688 | 0.010614 |
| ARHGAP33 | 3     | 1.861599 | 0.03963 |
| NASP  | 3     | 3.513297 | 0.014856 |
| FBKP9L | 3     | 2.481728 | 0.000524 |
| PPAT  | 3     | 5.720153 | 0.11E-06 |
| DCAF1L2 | 3     | 0.545007 | 0.036998 |

**From Table S4: Del-CIN facilitator**

| gene  | table | hr    | p value |
|-------|-------|-------|---------|
| POUAF1 | 4     | 2.743688 | 0.010614 |
| PRR19 | 4     | 2.005502 | 0.048102 |
| FBKP10 | 4     | 2.505569 | 0.029749 |
| PAICS | 4     | 4.351303 | 0.003473 |
| RAPAP3 | 4     | 2.938046 | 0.010397 |

**From Table S5: Amp-CIN suppressor**

| gene  | table | hr    | p value |
|-------|-------|-------|---------|
| IFFO1 | 5     | 1.981091 | 0.047718 |
| C16orf89 | 5     | 0.22141 | 0.039903 |
| SLC4A1A4 | 5     | 0.338891 | 0.02062 |
| MEF2C | 5     | 0.336933 | 0.044345 |
| GJA4 | 5     | 2.726681 | 0.030345 |
| KLHL2 | 5     | 0.124123 | 0.038676 |
| LCK | 5     | 2.524544 | 0.019527 |
| CD4 | 5     | 1.869079 | 0.030867 |
| CCL14 | 5     | 3.715572 | 0.010582 |
| SEPP1 | 5     | 0.464588 | 0.020495 |
| S1PR1 | 5     | 2.214425 | 0.005356 |
| GZMA | 5     | 2.585338 | 0.024011 |
| NEK5 | 5     | 2.212391 | 0.031957 |
| CCR1 | 5     | 4.863232 | 0.030661 |
| THEMIS | 5     | 2.745081 | 0.010459 |
| TCTEX1D1 | 5     | 2.517451 | 0.011728 |
| ITGAL | 5     | 0.336988 | 0.017752 |

**From Table S6: Del-CIN suppressor**

| gene  | table | hr    | p value |
|-------|-------|-------|---------|
| PRX | 6     | 2.356857 | 0.02928 |
| SLC4A1A4 | 6     | 0.338891 | 0.02062 |
| TREML1 | 6     | 0.32712 | 0.029162 |
| DNMBP | 6     | 2.517727 | 0.008073 |
| S1PR1 | 6     | 2.214425 | 0.003556 |
| NBEAL2 | 6     | 3.532039 | 0.000656 |
| C1orf158 | 6     | 3.415105 | 0.007369 |
| ZNF264 | 6     | 3.193597 | 0.003192 |
| ALOX5 | 6     | 0.200569 | 0.024066 |
| IFFO1 | 6     | 1.981091 | 0.047718 |
| MIA | 6     | 3.272932 | 0.045768 |

Red: HR>1. Blue: HR<1

(A)

Figure 4. Cont.
(B) Figure 4. Cont.
Figure 4. Cont.
Figure 4. CNA facilitator/suppressor affecting patients’ survival. (A) 39 genes for which expression levels correlate with both CNA and survival. Genes are shown indicating which category/Supplementary Table they are from. Highlighted in Red: HR > 1 (for which expression alterations increase risk). Blue: HR < 1. (B) PPAT overexpression (p = 0.0098) decreases survival (HR = 5.7, p = 3.11 × 10−6). PPAT (Phosphoribosyl Pyrophosphate Amidotransferase) catalyzes the first step of de novo purine nucleotide biosynthetic pathway. (C) PAICS overexpression (p = 0.0032) decreases survival (HR = 4.6, p = 0.0034). PAICS (Phosphoribosylaminomimidazole Carboxylase and Phosphoribosylaminomimidazolesuccinocarboxamidase Synthase) belongs to the same purine biosynthetic pathway and catalyzing steps six and seven. (D) From many immune components, survival-critical 6 genes are enriched in CD4-LCK-MEF2C axis and CCL14-CCR1 axis, thus these two axes emerge as survival-critical pathways. Pathway analysis illuminates CD4-LCK-MEF2C axis and CCL14-CCR1 axis as critical pathways for patients’ survival. These two pathways may represent critical immunomodulation targets. (E) Decreased expression of KLHL2 (p = 6.95 × 10−4) is a marker for better prognosis (HR = 0.1, p = 0.0387). (F) Increased expression of PRX (Periaxin) (p = 0.0017) is a marker for poor prognosis (HR = 2.3, p = 0.0293). (G) ALOX5-mediated leukotriene biosynthesis plays a role in survival (HR = 0.2, p = 0.0241).
4. Discussion

Previous works on cancer immunity or transcriptome had limited sample sizes and/or tested specific genes/pathways, e.g., [33,34]. The present study utilized 560 pooled non-overlapping lung adenocarcinoma data in the TCGA genome database and tested all 20,000+ human genes, and thus illustrates a GE landscape in an unbiased manner. This “gene expression to CNA (GE-CNA)” data-mining strategy has comprehensively uncovered 1578 genes associated with human lung adenocarcinoma CNA and genomic stability, and the pathways to which they belong. 492 CNA facilitator genes (Supplementary Table S1; Supplementary Figure S2A) and 1086 CNA suppressor genes (Supplementary Table S2; Supplementary Figure S2B) indicated a stark contrast in their functions, rather than simple loss- or gain-of-function. At the same time, the GE-CNA strategy results showed some similarities with the results of previous CNA-GE analysis [14].

To obtain additional mechanistic insights, we subcategorized the identified genes into four groups: 163 amplification/insertion CNA facilitator genes (Supplementary Table S3; Figure 2A); 187 deletion CNA facilitator genes (Supplementary Table S4; Figure 2B); 420 amplification/insertion CNA suppressor genes (Supplementary Table S5; Figure 3A); and 407 deletion CNA suppressor genes (Supplementary Table S6; Figure 3B). In theory, amplification/insertion CNA facilitator genes (Supplementary Table S3; Figure 2A) and amplification/insertion CNA suppressor genes (Supplementary Table S5; Figure 3A), and deletion CNA facilitator genes (Supplementary Table S4; Figure 2B) and deletion CNA suppressor genes (Supplementary Table S6; Figure 3B), are functionally antagonistic. The antagonistic relationship suggests that (a) amplification/insertion CNA associated with MIN or CIN with DNA replication stress can be suppressed by a broad range of immune cells (T, B, NK-cells, leukocytes), while (b) deletion CNA associated with mitotic regulator overexpression can be suppressed more specifically by leukocytes guided by leukocyte extravasation signaling. This prediction directs our attention to leukocytes in suppressing mitotic error-mediated “deletion CIN” tumor cells for designing novel modalities of immunotherapy in the future. CIN cancers respond to PD1/PD-L1 blockade immunotherapies poorly compared with MIN cancers [35], thus PD1/PD-L1 blockade therapies are approved for MIN cancers by the US FDA [36]. This phenomenon may be in part caused by the diversification of the immune pathways involved in targeting CIN and MIN.

Our discoveries are consistent with the notion that chromosome mis-segregation generates senescent cells with complex karyotypes/aneuploids and that they are eliminated by the immune system [15]. Yet, previous studies mainly reported innate characteristics of aneuploid cells that elicit immune activation, and only a limited number of immune cells or the trigger that responds to the activation signaling have been identified. For example, poly/hyperploid-type aneuploidy is sensed by immune cells with ER chaperone calreticulin, then eliminated by immunogenic death. Cells expressing NKG2D and DNAM1 ligands can be selectively targeted by NK cells [37,38]. This GE-CNA data-mining project successfully identified these pathways (Natural Killer Cell Signaling). Indeed, a majority of the CNA suppressors are immune signatures. The importance of immunomodulation in cancer therapy has been well established. Intervening PD1/PDL1 and CTLA4 pathways emerged as cancer immunotherapy modalities with efficacy and are being actively pursued [39]. Both pathways were identified (CTLA4 Signaling in Cytotoxic T Lymphocytes and PD-1, PD-L1 cancer immunotherapy pathway), as additional validation for this screening. Consistent with identification of the CD4-LCK-MEC2C axis, CD4 was identified as a key contributor to the PD1/PD-L1 immunotherapy efficacy, showing a critical role of CD4-mediated pathway in antagonizing cancer [40].

Newly illuminated with the present study and of particular clinical interest are survival-critical 39 genes and pathways of purine biosynthesis (PPAT, PAICS), immune-regulating CD4-LCK-MEC2C and CCL14-CCR1 axes, E3 ubiquitin ligase (KLHL2), PRX, and arachidonic acid metabolism ALOX5. In other organs (e.g., colon, pancreas), ALOX5 inhibition with COX1/2 inhibitors shows cancer chemoprevention effects [41–43]. In lung, ALOX5 may play a surveillance role for cells with CNA, as other researchers have sug-
gested that ALOX5 inhibition can promote tumor development [44,45]. Overall, this study identified target genes and pathways for future cancer chemoprevention, therapy, and immunomodulatory approaches for improved patient survival.

This data-mining strategy may have missed some modalities of eliminating cells with genomic instability at an early phase, as the datasets used are for fully developed lung adenocarcinomas. For example, micronuclei-bearing mitotically-failed cells can activate the cGAS/STING pathway via cytosolic DNA [46–48], yet the cGAS/STING pathway did not appear. It can be presumed that the cGAS/STING pathway is involved in removal of very early-stage cancer or single micronuclei-bearing mitotically-failed cells, and the signaling may not be pronounced in lung adenocarcinoma in advanced stages. In the future, the use of datasets from early stage cancers or pre-cancerous lesions may provide information relevant to early cancer surveillance and prevention.

5. Conclusions

The GE-CNA strategy applied on human lung adenocarcinoma revealed 1578 genes whose expression level correlates with increased or decreased CNA. Pathway analysis on the genes suggests competitive and CNA type-specific relationships between CNA facilitator genes, mostly mitotic and cell cycle regulators, and suppressor genes, whose functions mainly involved in immune system thus likely consist of CNA immunosurveillance mechanisms. Secondary screening identified survival critical 39 genes among the 1578 genes. They represent potential targets for drug and/or immunomodulation approaches against lung adenocarcinoma.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cancers13112586/s1, Figure S1: Results from a prototype/proof-of-principle study (AACR2019 Abstract #1199) Figure S2: CNA facilitators and CNA suppressors Table S1: CNA facilitator genes. Table S2: CNA suppressor genes. Table S3: Insertion/amplification CNA facilitator genes. Table S4: Deletion CNA facilitator genes. Table S5: Insertion/amplification CNA suppressor genes. Table S6: Deletion CNA suppressor genes. Table S7: Survival critical genes.

Author Contributions: Conception and design: H.Y.Y. Development of methodology: H.Y.Y., M.F., C.X. Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H.Y.Y., M.F., C.V.R., A.S.A., Y.Z., C.X. Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): C.V.R., H.Y.Y., M.F., C.X. Writing, review, and/or revision of the manuscript: C.V.R., C.X., H.Y.Y. Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H.Y.Y., M.F., C.X., C.V.R. Study supervision: H.Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact. This work was supported by the Carley-Cade chair fund (OUHSC) to C.V.R.; and the research support fund (Stephenson Cancer Center) and an Oklahoma Nathan Shock Center pilot study grant (OUHSC) to H.Y.Y.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: A public database (cbioportal.org) was utilized.

Acknowledgments: We would like to thank Elizabeth Cambron and Taylor McCoy for administrative assistance, Kathy Kyler for editorial aid, and members of the Center for Cancer Prevention and Drug Development for insightful discussions.

Conflicts of Interest: The authors declare no potential conflicts of interest.

References

1. McGranahan, N.; Burrell, R.A.; Endesfelder, D.; Novelli, M.R.; Swanton, C. Cancer chromosomal instability: Therapeutic and diagnostic challenges. *EMBO Rep.* 2012, 13, 528–538. [PubMed]
2. Walther, A.; Houlston, R.; Tomlinson, I. Association between chromosomal instability and prognosis in colorectal cancer: A meta-analysis. Gut 2008, 57, 941–950. [CrossRef] [PubMed]

3. Zhang, W.; Mao, J.-H.; Zhu, W.; Jain, A.K.; Liu, K.; Brown, J.B.; Karpen, G.H. Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. Nat. Commun. 2016, 7, 12619. [CrossRef]

4. McGranahan, N.; Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell 2017, 168, 613–628. [CrossRef]

5. Michel, L.S.; Liberal, V.; Chatterjee, A.; Kirchwegger, R.; Pasche, B.; Gerald, W.; Dobre, M.; Sorger, P.K.; Murty, V.V.V.S.; Benezra, R. MAD2 haplo-insufficiency causes premature anaphase and chromosome instability in mammalian cells. Nat. Cell Biol. 2001, 409, 355–359. [CrossRef]

6. Yamada, H.Y.; Zhang, Y.; Reddy, A.; Mohammed, A.; Lightfoot, S.; Dai, W.; Rao, C.V. Tumor-promoting/progressing role of additional chromosome instability in hepatic carcinogenesis in Sgo1 (Shugoshin 1) haploinsufficient mice. Carcinogenesis 2015, 36, 429–440. [CrossRef]

7. Yamada, H.Y.; Kumar, G.; Zhang, Y.; Rubin, E.; Lightfoot, S.; Dai, W.; Rao, C.V. Systemic chromosome instability in Shugoshin-1 mice resulted in compromised glutathione pathway, activation of Wnt signaling and defects in immune system in the lung. Oncogenesis 2016, 5, e256. [CrossRef]

8. Sotillo, R.; Schwartzman, J.M.; Socci, N.D.; Benezra, R. Mad2-induced chromosome instability leads to lung tumour relapse after on-cogene withdrawal. Nature 2010, 464, 436–440. [CrossRef]

9. Simon, J.E.; Bakker, B.; Foijer, F. CINcere Modelling: What Have Mouse Models for Chromosome Instability Taught Us? Recent Results Cancer Res. 2015, 200, 39–60.

10. Bakhoun, S.F.; Compton, D. Chromosomal instability and cancer: A complex relationship with therapeutic potential. J. Clin. Investig. 2012, 122, 1138–1143. [CrossRef] [PubMed]

11. Rao, C.V.; Yamada, H.Y. Genomic Instability and Colon Carcinogenesis: From the Perspective of Genes. Front. Oncol. 2013, 3, 130. [CrossRef] [PubMed]

12. Kops, G.J.P.L.; Weaver, B.A.A.; Cleveland, D.W. On the road to cancer: Aneuploidy and the mitotic checkpoint. Nat. Rev. Cancer 2005, 5, 773–785. [CrossRef]

13. Huang, H.; Yen, T.J. Mitotic Checkpoint and Chromosome Instability in Cancer. In Cell Cycle Deregulation in Cancer; Enders, G., Ed.; Springer: New York, NY, USA, 2009; p. 19.

14. Davoli, T.; Uno, H.; Wooten, E.C.; Elledge, S.J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. Science 2017, 355, eaaf8399. [CrossRef] [PubMed]

15. Santaguida, S.; Richardson, A.; Iyer, D.R.; M’Saad, O.; Zasadil, L.; Knouse, K.A.; Wong, Y.L.; Rhind, N.; Desai, A.; Amon, A. Chromosome Mis-segregation Generates Cell-Cycle-Arrested Cells with Complex Karyotypes that Are Eliminated by the Immune System. Dev. Cell 2017, 41, 638–651.e5. [CrossRef] [PubMed]

16. Gao, J.; Aksoy, B.A.; Dogrusoz, U.; Dresdner, G.; Gross, B.; Sumer, S.O.; Sun, Y.; Jacobsen, A.; Sinha, R.; Larsson, E.; et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. Sci. Signal. 2013, 6, pl1. [CrossRef]

17. Cerami, E.; Gao, J.; Dogrusoz, U.; Gross, B.E.; Sumer, S.O.; Aksoy, B.A.; Jacobsen, A.; Byrne, C.J.; Heuer, M.L.; Larsson, E.; et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discov. 2012, 2, 401–404. [CrossRef] [PubMed]

18. Li, B.; Ruotti, V.; Stewart, R.M.; Thomson, J.A.; Dewey, C.N. RNA-Seq gene expression estimation with read mapping uncertainty. Bioinformatics 2010, 26, 493–500. [CrossRef]

19. Merkel, C.H.; Schumacher, S.E.; Hill, B.; Meyerson, M.L.; Beroukhim, R.; Getz, G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011, 12, 1–14. [CrossRef]

20. Storey, J.D.; Bass, A.J.; Dabney, A.; Robinson, D. Q-Value: Q-Value Estimation for False Discovery Rate Control. R Package Version 2.2.0. Available online: http://github.com/jdstorey/qvalue (accessed on 1 May 2020).

21. Krämer, A.; Green, J.; Pollard, J.; Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics 2014, 30, 523–530. [CrossRef]

22. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple hypothesis test-ing. J. R. Stat. Soc. B 1995, 57, 289–300.

23. Wilhelm, T.; Said, M.; Naim, V. DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. Genes 2020, 11, 642. [CrossRef] [PubMed]

24. Han, Y.; Wang, X. The emerging roles of KPNA2 in cancer. Life Sci. 2020, 241, 117140. [CrossRef]

25. Kimura, T.; Cui, D.; Kawano, H.; Yoshitomi-Sakamoto, C.; Takakura, N.; Ikeda, E. Induced expression of GINS complex is an essential step for reactivation of quiescent stem-like tumor cells within the peri-necrotic niche in human glioblastoma. J. Cancer Res. Clin. Oncol. 2015, 141, 365–371. [CrossRef]

26. Bondarieva, A.; Raveendran, K.; Telychko, V.; Rao, H.B.D.P.; Ravindranathan, R.; Zorzompokou, C.; Finsterbusch, F.; Dereli, I.; Papanikos, F.; Tränkner, D.; et al. Proline-rich protein PRR19 functions with cyclin-like CNTD1 to promote mitotic crossing over in mouse. Nat. Commun. 2020, 11, 1–20. [CrossRef] [PubMed]

27. Zhang, L.; Xu, H.; Ma, C.; Zhang, J.; Zhao, Y.; Yang, X.; Wang, S.; Li, D. Upregulation of deubiquitinase PSMD14 in lung adenocarcinoma (LUAD) and its prognostic significance. J. Cancer 2020, 11, 2962–2971. [CrossRef] [PubMed]
28. Tanenbaum, M.E.; Macurek, L.; van der Vaart, B.; Galli, M.; Akhmanova, A.; Medema, R.H. A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by Aurora kinases. *Curr. Biol.* 2011, 21, 1356–1365. [CrossRef]

29. Famulski, J.K.; Vos, L.; Sun, X.; Chan, G. Stable hZW10 kinetochore residency, mediated by hZwint-1 interaction, is essential for the mitotic checkpoint. *J. Cell Biol.* 2008, 180, 507–520. [CrossRef] [PubMed]

30. Jiang, H. The complex activities of the SET1/MLL complex core subunits in development and disease. *Biochim. Biophys. Acta Bioenerg.* 2020, 1863, 194560. [CrossRef] [PubMed]

31. Wettachureck, N.; Strilic, B.; Offermanns, S. Passing the Vascular Barrier: Endothelial Signaling Processes Controlling Extravasation. *Physiol. Rev.* 2019, 99, 1467–1525. [CrossRef]

32. Goswami, M.T.; Chen, G.; Chakravarthi, B.V.; Pathi, S.S.; Anand, S.K.; Carskadon, S.L.; Giordano, T.J.; Chinnaiyan, A.M.; Thomas, D.G.; Palanisamy, N.; et al. Role and regulation of coordinately expressed de novo purine biosynthetic en-zymes PPAT and PAICS in lung cancer. *Onco target* 2015, 6, 23445–23461. [CrossRef] [PubMed]

33. Naour, J.L.; Zitvogel, L.; Galluzzi, L.; Vacchelli, E.; Kroemer, G. Trial watch: STING agonists in cancer therapy. [CrossRef] [PubMed]

34. Tijhuis, A.E.; Johnson, S.C.; McClelland, S.E. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. *Mol. Cytogenet.* 2019, 12, 1–21. [CrossRef]

35. Tijhuis, A.E.; Johnson, S.C.; McClelland, S.E. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. *Mol. Cytogenet.* 2019, 12, 1–21. [CrossRef]

36. Naour, J.L.; Zitvogel, L.; Galluzzi, L.; Vacchelli, E.; Kroemer, G. Trial watch: STING agonists in cancer therapy. *Oncoimmunology* 2020, 9, 1777624. [CrossRef]