Journal Pre-proof

Exploring the molecular landscape of NNK-induced transformation: A comprehensive genome-wide CRISPR/Cas9 screening

Trang Dinh, Mira Rahm, Zhenghe Wang, Christopher McFarland, Athar Khalil

Please cite this article as: Dinh T, Rahm M, Wang Z, McFarland C, Khalil A, Exploring the molecular landscape of NNK-induced transformation: A comprehensive genome-wide CRISPR/Cas9 screening, *Genes & Diseases*, https://doi.org/10.1016/j.gendis.2023.101131.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Exploring the molecular landscape of NNK-induced transformation: A comprehensive genome-wide

CRISPR/Cas9 screening

Lung cancer is the leading cause of cancer-related deaths worldwide, with an estimated 1.8 million deaths in 2020 alone [1]. While several risk factors are associated with lung cancer, smoking remains the most significant environmental cause of the disease [1]. One of the most potent carcinogens in tobacco smoke is 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK), which is produced from nicotine through a series of metabolic reactions [1,2]. NNK has been shown to cause DNA damage, activate oncogenic signaling pathways, and promote cell proliferation and survival, all of which are hallmarks of cancer development [1]. Despite extensive research on the molecular mechanism of action of NNK, direct targets that mediate sensitivity to NNK and contribute to its tumorigenicity remain unknown.

Various experimental approaches, including colony formation assays, proliferation studies, and tumor growth in nude mice, have consistently demonstrated NNK facilitation of malignant transformation in human bronchial epithelial cells [1]. The BEAS-2B cell line is an immortalized but non-tumorigenic line derived from human bronchial epithelium, ideal for understanding this malignant transformation. This cell line has been widely used as a model system to study lung carcinogenesis [1]. In our study, we comprehensively elucidated the molecular regulatory targets underlying NNK's effect on human lung cancer progression through CRISPR/Cas9 genome-wide knockout screening in BEAS-2B cells. We identified key cellular pathways associated with sensitivity and resistance to NNK, providing novel insights into the molecular mechanisms underlying NNK-induced tumorigenicity and offering potential targets for lung cancer prevention and treatment in smokers.

To gain a comprehensive understanding of potential molecular targets that mediate the sensitivity and resistance to acute and chronic NNK exposure, BRAS-2B cells were transduced with the Brunello CRISPR genome-wide knockout pooled library (Catalog # 73179-LV, 1-vector-system) consisting of 76,411 single-guide RNAs (sgRNAs) targeting 19,114 genes. Building upon the previously established data, we subjected transduced BEAS-2B cells to NNK treatment at a concentration of 100 mg/L for 24 h (Fig. 1A) [1]. Cell viability assays confirmed that NNK accelerates malignant transformation as evidenced by a significantly higher proliferation rate in Brunello-transduced BEAS-2B cells following NNK exposure as compared with the control-treated cells at two time points (1 and 4 weeks) (Fig. 1B). In both NNK-exposed and control lines, proliferation rate increased with time, consistent with an adaptive evolutionary process. Remarkably, after 72 hours of NNK treatment, visible colonies were detected in the transduced NNK-exposed cell population (Fig. S1). While maintaining a representation of at least 300 cells per sgRNA in each collection, we isolated gDNA from cells, and sgRNA genes were amplified by PCR and deep-sequenced at an average depth of 31 million reads per sample. After normalizing beta scores by cell cycle-relevant genes, we identified genes that modulate NNK sensitivity and resistance by computational analysis using the MAGeCK pipeline (Fig. S2).

To capture the influence of exposure time on NNK-related gene perturbations, we analyzed a nine-square scatterplot for early and late time points independently (Fig. 1C, D). We focused specifically on genes of which perturbations showed increased resistance or sensitivity to NNK treatment. Our analysis revealed 675 genes to be strongly enriched in the NNK sample compared with the control at the early time point (Fig. 1C), while 529 genes showed strong enrichment at the late time point (Fig. 1D). These genes likely play a role in conferring resistance to NNK-induced transformation, as their knockout sensitized

40   BEAS-2B cells to NNK treatment (Fig. S3A). On the flip side, our analysis revealed 142 genes that were significantly depleted

41   in the early time point, along with 24 genes in the late time point, suggesting their sensitivity and selectivity to NNK-mediated

42   transformation (Fig. S3B). Further investigation of commonly enriched and depleted genes at the two time points was

43   conducted to gain deeper insights into the molecular mechanisms underlying NNK-mediated effects (Fig. S4A, B). The 404

44   commonly enriched genes in our screen were also evaluated for overrepresentation of functional pathways using MAGeCK

45   (Fig. 1E). These genes encode cellular components intimately connected to cancer progression, including DNA replication,

46   cancer regulation by PD-L1, cellular senescence, cell cycle arrest, and signal regulation by P53, FoxO, and cGMP-PGK

47   signaling pathways. Thus, activation of these pathways and intact function of the associated genes may enhance DNA repair

48   mechanisms, immune surveillance, growth inhibitory responses, and stress-induced cell fate decisions, collectively reducing

49   the likelihood of NNK-mediated cellular transformation. Annotating the enriched genes in the identified pathways based on the

50   Cancer Gene Census database, we identified 13 hits among these commonly enriched genes as known tumor suppressors

51   (PRDM1, SETD7, ATM, RASSF5, TSC1, EXT1, PTPN1, SFRP4, CDKN2B, B2M, ARID1B, MRE11A, and PRKAG2) (Fig.

52   1G). Herein, the knockout of cyclin-dependent kinase 2 (CDKN2B) and histone-lysine N-methyltransferase (SETD7), both

53   well-known for their crucial roles in regulating the cell cycle, rendered BEAS-2B cells more vulnerable to the transforming

54   effects of NNK, potentially playing a significant role in protecting against the development of lung cancer upon smoking

55   exposure [3,4]. Additionally, the loss of genes such as the serine/threonine kinase ATM (ataxia telangiectasia mutated) and

56   meiotic recombination 11 homolog A (MRE11A) impaired the ability of the cells to properly detect and repair DNA damage

57   caused by NNK. This finding aligns with previous research demonstrating that reduced promoter activity of the MRE11A

58   gene, involved in recognizing DNA damage and activating ATM, increased the risk of lung cancer among smokers [4].

59   Interestingly, among our identified relevant tumor suppressors, mutations in two of them (CDKN2B and B2M) displayed a

60   worse survival rate among 466 lung cancer smoking patients across two lung adenocarcinoma studies (Fig. 1I). This

61   observation highlights the clinical relevance of these genes in the context of smoking-related lung cancer. Finally, the robust

62   enrichment of genes associated with cellular senescence highlights its crucial role in countering the harmful effects of NNK

63   and underscores its potential as a target for therapeutic interventions aimed at reducing NNK-induced lung carcinogenesis.

64

65   On the other side, the commonly depleted genes encode for pathways involved in Wnt and Hippo signaling as well as lipid

66   metabolism, all of which have been implicated in lung cancer development and tumor initiation and progression (Fig. 1F). In

67   our screen, the knockout of the underlying 79 commonly depleted genes in NNK samples inhibited the NNK transformation in

68   BEAS-2B cells indicating a potential sensitivity for NNK transforming effect of these genes (Fig. 1H). Notably, one of the

69   depleted sgRNA targets in NNK samples was cytochrome P450 family 4 subfamily A member 22 (CYP4A22), a member of

70   the cytochrome P450 (CYP) enzyme family [1]. Interestingly, CYP4A22 is an orphan CYP with an unknown function in the

71   context of NNK exposure. Knowing that the p450 cytochrome family is responsible for NNK metabolism, further investigation

72   is warranted to elucidate CYP4A22's precise role and mechanisms of action in NNK-induced effects [5].

73

74   Our study is the first to unveil a comprehensive landscape of genetic perturbations and potential targets involved in the

75   sensitivity and resistance to NNK, shedding light on unexplored mechanisms and targets for NNK-induced effects, while also

76   re-affirming that NNK directly interacts with DNA repair and cell cycle control machinery. Further investigations into these

77   identified pathways and genes will deepen our understanding of the molecular events driving NNK-induced transformation and

78 may contribute to the development of novel strategies for intervention, prevention, and personalized treatment of NNK- and

79 smoking-associated lung cancer.

80

84

85 **References**

86 1. Zhou J, Xiong R, Zhou J, et al. Involvement of m6A regulatory factor IGF$_2$BP$_1$ in malignant transformation of human

87 bronchial epithelial Beas-2B cells induced by tobacco carcinogen NNK. *Toxicol Appl Pharmacol*. 2022;436:115849.

88 2. Leng S, Stidley CA, Willink R, et al. Double-strand break damage and associated DNA repair genes predispose smokers to

89 gene methylation. *Cancer Res*. 2008;68(8):3049-3056.

90 3. Xia Y, Liu Y, Yang C, et al. Dominant role of CDKN2B/p15INK4B of 9p21.3 tumor suppressor hub in inhibition of cell-

91 cycle and glycolysis. *Nat Commun*. 2021;12(1):2047.

92 4. Batista IAA, Helguero LA. Biological processes and signal transduction pathways regulated by the protein

93 methyltransferase SETD7 and their significance in cancer. *Signal Transduct Target Ther*. 2018;3:19.

94 5. Durairaj P, Fan L, Machalz D, Wolber G, Bureik M. Functional characterization and mechanistic modeling of the human

95 cytochrome P450 enzyme CYP4A22. *FEBS Lett*. 2019;593(16):2214-2225.

96

97 Trang Dinh, Mira Rahm, Zhenghe Wang, Christopher McFarland[*], Athar Khalil[*]

98

99 *Department of Genetics and Genome Sciences, Case Western Reserve University, Cleveland 44106, United States.*

100

101 [*] Corresponding authors.

102 *E-mail addresses*: axk1472@case.edu (A. Khalil), cdm113@case.edu (C. McFarland).

103

104 **Figure 1** CRISPR/Cas9 genome-wide disruption screening in NNK exposed BEAS-2B cells and bioinformatic analyses. **(A)**

105

106 by cell counting during each collection and divided by seeding cell number (t-test, $P < 0.05$). **(C, D)** Nine-square scatterplots

107 for identifying treatment-associated genes. Genes in green and blue groups are found to be strongly negatively or positively

108 selected, respectively, in the control samples, but they are weakly selected in the treatment samples. Genes in the orange group

109 have a strong positive selection in NNK-treated cells as compared with control; these genes are denoted as NNK-resistant

110 genes. Purple-encoded genes are highly negatively selected in the treatment sample; these genes are denoted as NNK-sensitive

111 genes. The numbers in red indicate the number of genes classified in each group. **(E, F)** Functional analysis for treatment-

112 associated genes. Pathway enrichment analysis was done via MAGeCK for commonly enriched **(E)** and commonly depleted

113 **(F)** genes in NNK treatment identified by our CRISPR screening. Top-row dots represent significantly enriched pathways with

114 the most numbers of genes; dot size indicates the number of genes belonging to the noted pathway. Bottom-row dots represent

115 genes that belong to the above significantly enriched pathways; the dot color indicates the differential beta score of the noted

116 gene in NNK treatment versus control/methanol treatment. **(G, H)** Heatmap of the top commonly enriched NNK resistant

117 genes **(G)** and top commonly depleted NNK sensitive genes **(H)** among the identified pathways. Highlighted in green are 13

118 tumor suppressor genes (TSGs), annotated using the Cancer Gene Census dataset (Genome Version GRCh38 · COSMIC v98).

119 **(I)** Kaplan–Meier survival plot comparing relapse-free survival (RFS) of 466 smoking lung cancer patients with mutated (red

120 line) versus wild-type (blue line) CDKN2B ($P = 0.0266$) or $P = 2.089$e-3 genes. Significance was tested by log-rank test. Deep

121 deletion corresponded to 95.5% of CDKN2B mutations and truncating mutations corresponded to 57.1% of those in B2M. The

122 data were retrieved from the lung adenocarcinoma dataset of the Memorial Sloan Kettering Cancer Center study (MSK, J

123 Thorac Oncol 2020).

124