HHS Public Access

Author manuscript

Cancer Res. Author manuscript; available in PMC 2018 November 15.

Genomic landscape of atypical adenomatous hyperplasia reveals divergent modes to lung adenocarcinoma

Smruthy Sivakumar[1,2], F. Anthony San Lucas[1], Tina L. McDowell[3], Wenhua Lang[3], Li Xu[3], Junya Fujimoto[3], Jianjun Zhang[4], P. Andrew Futreal[5], Junya Fukuoka[6], Yasushi Yatabe[7], Steven M. Dubinett[8], Avrum E. Spira[9], Jerry Fowler[1], Ernest T. Hawk[10], Ignacio I. Wistuba[3], Paul Scheet[1,2,*], and Humam Kadara[1,11,*]

[1]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[2]The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences, Houston, Texas, USA

[3]Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[4]Department of Thoracic/Head and Neck Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[5]Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[6]Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, Japan

[7]Department of Pathology and Molecular Diagnostics, Aichi Cancer Center, Nagoya, Japan

[8]Division of Pulmonology and Critical Care Medicine, University of California Los Angeles, Los Angeles, California, USA

[9]School of Medicine, Boston University, Boston, Massachusetts, USA

[10]Division of Cancer Prevention, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[11]Department of Biochemistry and Molecular Genetics, The American University of Beirut, Beirut, Lebanon

Abstract

There is a dearth of knowledge about the pathogenesis of premalignant lung lesions, especially for atypical adenomatous hyperplasia (AAH), the only known precursor for the major lung cancer subtype adenocarcinoma (LUAD). In this study, we performed deep DNA and RNA sequencing

Correspondence: Paul Scheet, PhD., Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA, email: pscheet@alum.wustl.edu or to Humam Kadara, PhD., Department of Biochemistry and Molecular Genetics, Faculty of Medicine, American University of Beirut, Beirut, Lebanon; Department of Epidemiology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA, Tel: +9611350000; hk94@aub.edu.lb.
*Equally contributing co-corresponding authors

analyses of a set of AAH, LUAD and normal tissues. Somatic *BRAF* variants were found in 5/22 (23%) of AAH patients, 4/5 of whom had matched LUAD with driver *EGFR* mutations. *KRAS* mutations were present in all ever-smoker cases in the cohort (18%) exclusive of the cases with *BRAF* mutations. Integrative analysis revealed profiles expressed in *KRAS*-mutant cases (*UBE2C*, *REL*) and *BRAF*- mutant cases (*MAX*) of AAH, or common to both sets of cases (suppressed *AXL*). Gene sets associated with suppressed anti-tumor (Th1; *IL12A*, *GZMB*) and elevated pro-tumor (*CCR2*, *CTLA-4*) immune signaling were enriched in AAH development and progression. Our results reveal potentially divergent *BRAF* or *KRAS* pathways of AAH and immune dysregulation in the pathogenesis of this pre-malignant lung lesion.

Keywords

Atypical adenomatous hyperplasia; lung adenocarcinoma; pathogenesis; deep sequencing; immune signaling

INTRODUCTION

Non-small cell lung cancer (NSCLC) constitutes the majority (~85%) of lung malignancies [1]. NSCLC is also the leading cause of cancer mortality [2]. This is largely attributed to late diagnosis after locoregional or distant spread [1]. Advances (e.g., molecular-based) in early detection and prevention of NSCLC have been limited by a poor understanding of early changes in NSCLC pathogenesis [3].

NSCLC is mainly comprised of two histological subtypes, adenocarcinomas (LUADs) and squamous cell carcinomas (LUSCs), with LUAD the most common of these [1]. Studies have demonstrated that LUADs and LUSCs carry different genomic landscapes of driver mutations as well as disparate clinicopathological features [3]. While LUSCs are diagnosed almost exclusively in smokers, LUADs are known to develop in both smokers and non-smokers [3]. In contrast to the well characterized histopathological sequence of various lesions (bronchial hyperplasias, metaplasias and dysplasias, as well as carcinomas *in situ*) associated with LUSC pathogenesis, atypical adenomatous hyperplasias (AAHs) are the only known precursor (premalignant) lesions in the development of LUAD [3,4].

Very few molecular alterations have been described in AAHs. Mutations in *EGFR* and *KRAS* have been identified previously in AAH, showing mutual exclusivity [5]. Whereas, *EGFR* mutations were shown to present at similar frequencies in AAH and LUAD, *KRAS* mutations (e.g., codon 12), were demonstrated to occur more frequently in AAH compared to LUAD [5]. Further, it has also been shown that *KRAS*-mutant AAHs are more prevalent in smokers while *EGFR*-mutant AAHs do not show association with smoking status [5]. Loss-of-heterozygosity of chromosome arms 3p, 9p, 16p, 17q, and 17p have been shown in AAH, and in common with corresponding primary LUADs [3]. Other molecular aberrations that have been identified in AAH include overexpression of Cyclin D1, survivin, *ERBB2* oncoproteins and *NKX2-1* [3]. Epigenetic modifications such as increased promoter hypermethylation of genes implicated in lung cancer (e.g., *p16*) with histologic progression from normal to AAH and finally LUAD [6], as well as DNA methylation of CDKN2A and PTPRN2, have also been reported in AAHs [7]. A recent multi-region sequencing study

identified several clonal events predicted to be involved in the transformation of these precursor lesions and that were further shown to be present in paired circulating DNA [8]. These studies point to the molecular complexity of AAH, which is still poorly understood, and its role as a precursor to LUAD.

We interrogated, by deep sequencing, somatic mutations and gene expression changes in normal tissues, AAHs and malignant tumors from patients with LUAD. We identify subgroups of AAHs with distinct driver mutations, expression signatures and aberrant markers of the immune response that inform the molecular pathogenesis of AAH. These alterations offer a window to explore new strategies for early clinical management (e.g. detection and targeted chemoprevention) of LUAD.

MATERIALS AND METHODS

Cohort of cases with atypical adenomatous hyperplasias

Normal lung parenchyma tissues (NL), atypical adenomatous hyperplasias (AAHs) and lung adenocarcinomas (LUADs) ($n = 67$ samples) were acquired from 22 patients with early-stage LUAD who were evaluated at the Aichi Cancer Center (Nagoya, Japan) and Nagasaki University (Nagasaki, Japan). Specimens were approved for study by institutional review boards and according to the international ethical guidelines for biomedical research involving human subjects (CIOMS). Informed written consents were received from all subjects wherever necessary. Cases with available AAHs were included in the analysis to study mutational and expression signatures involved in the pathogenesis of AAH (Figure 1). Clinicopathological features of these patients are summarized in Table 1. The diagnosis, specimen collection and slide preparation were carried out between 2011 and 2015 for all patients. Specimens were obtained formalin-fixed and paraffin-embedded (FFPE) and stained by hematoxylin and eosin/H&E. A tabulation of cases and samples analyzed by DNA- (22 cases) and RNA- (17 cases) sequencing is provided in Table S1.

DNA and total RNA isolation

DNA/RNA was extracted from all samples using the AllPrep DNA/RNA FFPE kit from Qiagen and suspended in nuclease free water (RNA) or AE buffer (DNA). Sample concentrations were measured on a NanoDrop 1000 (Thermo Fisher Scientific), and RNA integrity numbers indicative of overall quality were obtained on the 2100 Bioanalyzer (Agilent Technologies) using the RNA 6000 Nano or Pico kit based on RNA concentrations and according to the manufacturer's protocol. DNA was quantified using the Quant-iT PicoGreen double stranded DNA (dsDNA) kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Deep targeted DNA sequencing

The Ion AmpliSeq™ Comprehensive Cancer Panel (CCP; Thermo Fisher Scientific) comprising primers for 409 canonical cancer-associated genes and the AmpliSeq Library Kit 2.0 (Thermo Fisher Scientific) were used to prepare barcoded libraries from the FFPE DNA samples. Sequencing was performed on the Ion Torrent Proton platform according to the

manufacturer's instructions. The targeted DNA sequencing data files have been deposited in the sequence read archive (SRA) under Bioproject accession PRJNA398260.

Identification of somatic mutations

Using BAM files generated in the Ion Torrent server, somatic variants were rigorously identified based on four different variant calling methods: the Ion Torrent proprietary software Ion Reporter, marginal VCFs generated from Torrent Variant Caller (TVC), MuTect [9] and Varscan2 [10]. Instead of using all variants detected by any of the four callers, for each sample, the stringency was increased to include variants that were detected by at least two different callers, with the exception of those identified by Ion Reporter software and TVC. Variants in exonic, splicing and untranslated regions (UTRs) were assessed, focusing on exonic single nucleotide variants (SNVs) within the targeted 409 cancer gene panel. Nonsynonymous mutations in genes considered to be *bona fide* drivers of cancer according to the previous report by Vogelstein *et al.* [11] were analyzed in AAHs and LUADs. Mutations in genes previously determined by the cancer genome atlas (TCGA) to be significantly mutated in LUAD [12] were also assessed in both AAHs and LUADs.

Transcriptome sequencing

A subset of the cases (15 NLs, 17 AAHs and 16 LUADs from 17 different cases) was selected for transcriptome sequencing based on specimen availability as well as on transcriptome sequencing quality indicated by percentage of mapped reads and valid on-target reads. All samples were reverse-transcribed to generate cDNA libraries using the Ion AmpliSeq Transcriptome Human Gene Expression Kit (Thermo Fisher Scientific) adhering to the manufacturer's protocol. For sequencing, specimens from two cases were processed together in one chip. All samples were sequenced on an Ion Proton sequencer. Raw transcriptome sequence data files have been deposited in the gene expression omnibus under data series GSE102511.

Expression analysis

Transcriptomes were quantified from BAM alignment files generated in the Ion Torrent server using a expectation-maximization (E/M) algorithm based procedure [13]. Resultant gene-based counts were normalized, log (base 2) transformed, and corrected for batch using the R limma package [14]. Hierarchical clustering using Pearson correlation was performed in R. Pathways, gene-network identification and gene set enrichment analyses were performed using the commercially available software Ingenuity Pathways Analysis (IPA).

Additional details are found in the Supplementary Methods accompanying the manuscript.

RESULTS

Molecular profiling of normal lung tissue, atypical adenomatous hyperplasia and early-stage lung adenocarcinoma

There is a limited understanding of molecular alterations in AAH, the only known precursor lesion for the major lung cancer subtype adenocarcinoma [3,4]. Here, we sought to survey mutation and expression profiles in the pathogenesis of AAHs by studying these

premalignant lesions along with normal tissues and primary LUADs (Figure 1) from 22 patients with early-stage LUAD (Table 1). Diagnosis and histopathological determination of specimens (normal lung, AAH, LUAD) were conducted by analysis of H&E staining of five micron sections (Figure 1) and according to the WHO report on lung tumor classification by Travis *et al* [15]. Among the 22 patients, 12 were ever-smokers and 10 were non-smokers (Table 1). The cohort comprised 12 females and 10 males with a median age of 64. The majority (*n* = 18) of LUADs were stage IA with the remaining four cases determined to be IB (*n* = 3) and IIIA (*n* = 1) (Table 1). We studied matched samples (*n* = 67; one patient comprised two tumor specimens; Table S1) from the 22 early-stage LUAD patients by deep (mean = 961×) targeted sequencing of a 409 pan-cancer gene panel (Table S2). A subset of the cases (*n* = 17) and samples (*n* = 48), based on availability and sequencing quality of AAHs, were also surveyed by transcriptome sequencing (Table S1). Details of the quality metrics for each sequencing platform are listed in Table S2.

Mutational landscape of atypical adenomatous hyperplasia

All 45 AAHs and LUADs exhibited at least one somatic variant (exonic, splicing or in UTRs) with a mean of 6.1 variants (min = 1, max = 19) in AAHs and 10.6 (min = 1, max = 60) in LUADs (Figure S1). Non-smokers displayed an expected lower somatic mutation burden than ever-smokers in the LUADs (5.4 vs 14.5); however, they exhibited a similar burden in AAHs (6.4 vs 5.9; Figure S2). We then specifically interrogated in AAH, nonsynonymous mutations in genes considered to be *bona fide* drivers of cancer [11] (Figure 2A). We also examined mutations in genes previously determined by TCGA to be significantly recurrently altered in LUAD [12]. AAHs from five patients (23%) exhibited somatic activating mutations in the *BRAF* oncogene (Figure 2A). Interestingly, the *BRAF* mutations were not detected in the paired LUAD specimens (Figure 2A; Figure 2B; Table S3; Table S4). Four of the five AAHs exhibited a *BRAF* p.K601E mutation, the other AAH contained a *BRAF* p.N581S variant (Figure 2C, Table S3). *KRAS* was the second most recurrently mutated gene in AAHs (four cases, 18%; Figure 2A). All four *KRAS*-mutant premalignant tissues were from ever-smokers, in contrast to *BRAF*-mutant AAHs (from three non-smokers and two ever-smokers; Figure 2A). Also, AAH *KRAS* and *BRAF* mutations showed mutual exclusivity (Figure 2A). An interesting observation was that for four of the five cases (80%) with *BRAF*-mutant AAHs, their paired LUADs harbored activating mutations (three p.L858R in exon 21 and one p.S752F in exon 19; Table S3) of the *EGFR* oncogene (Figure 2A). The other LUAD exhibited inactivating mutations in *KEAP1* and *STK11* tumor suppressors (Table S3). LUADs of cases with the *KRAS*-mutant AAHs exhibited mutations in other drivers besides *KRAS* such as *TP53* (Figure 2A). We further validated by digital PCR, the presence of all sequencing derived *BRAF* and *KRAS* mutations in AAHs, and the *EGFR* p.L858R mutation in both AAHs and LUADs that exhibited these patterns (Table S5; Figure S3). Variant allele frequencies (VAFs) based on digital PCR were consistent with sequencing-based VAFs for these loci (Table S5). *TSC1* was the most frequently mutated tumor suppressor in AAHs (13.6%; two nonsense and one missense mutation; Figure 2A; Table S3). We also noted other mutated oncogenes (*EGFR* and *JAK3*) and tumor suppressors (*CDKN2A* and *TP53*) in AAHs (Figure 2A). Additionally in this cohort, 28 genes were mutated in both AAHs and LUADs (e.g., *KRAS, TP53, KEAP1, CDKN2A*), 84 were found only in LUADs (*STK11, PIK3CA*) and 29 were found

only in the preneoplastic lesions (*AMER1, BRAF, KDM5C, ERBB2*) (Figure 2B; Table S4). Even for genes that were shared between AAH and LUAD tissues, there were notable examples of differential frequency (e.g., *KRAS* more common in AAH; *EGFR* and *TP53* more common in LUAD; Figure S4). On further examination of these 28 shared genes, we found that *EGFR* and *KAT6B* exhibited the same mutations in both tissues (Table S3). There was also an enrichment of different mutations on codon 12 of *KRAS* in both the tissues (Table S3). Our findings underscore subgroups of AAH with different mutated driver genes (*BRAF* vs *KRAS*) suggestive of potentially various mechanisms in the pathogenesis of these premalignant lesions.

Expression profiles in the development and progression of atypical adenomatous hyperplasia

Next we sought to characterize expression profiles signifying the development of AAH from normal lung tissue (NL), and its progression to LUAD. We performed RNA-sequencing of a subset of the cases and samples (15 NLs, 17 AAHs and 16 LUADs – from 17 patients), based on availability and quality of AAHs, using a capture method targeting over 20,000 Refseq genes. Using an ANOVA model with tissue type (NL, AAH and LUAD) as a factor and patient as a random effect, with a test of no expression difference ($P < 0.001$) and an observed two-fold change minimum among pairwise tissue comparisons, we identified 1,008 genes differentially expressed in at least one of three tissue types (Figure 3; Table S6). Using one-sided t-tests to interrogate the two-step (NL to AAH and AAH to LUAD) modes of differential expression, we identified eight patterns or clusters of expression among the identified profiles (Figure 3). These consisted of the following: decrease ($n = 214$) from NL to AAH and from AAH to LUAD; increase ($n = 204$) from NL to AAH and from AAH to LUAD; decrease ($n = 116$) and increase ($n = 146$) from NL to AAH alone (no change from AAH to LUAD); decrease ($n = 85$) and increase ($n = 126$) from AAH to LUAD alone (no change from NL to AAH) and, less prevalent forms with no net change in expression such as an increase ($n = 33$) or decrease ($n = 84$) in AAH alone relative to other tissues (Table S6). A pathway based enrichment analysis was performed for genes in each cluster to infer potentially altered signaling (Figure 3). This analysis pinpointed decreased anti-tumor T-helper (Th1) immunity, and conversely, increased pro-tumor Th 2-based immune response and signaling in both phases, the development of AAH from NL and their progression to LUAD. Inhibition of *IFN-γ* and *TGFB1* signaling occurred early in AAH, when compared to NLs, and reduced surfactant protein signaling occurred thereafter in LUAD only. Pathways and gene set enrichment analysis also revealed an activation of B-cell receptor, *CSF2* (indicative of pro-tumor immune response), *MYC* and *ERBB2* signaling in AAH and LUAD (Figure 3). Activation of *WNT* and β-catenin signaling as well as modulation of gene sets associated with increased immune cell (phagocytes) migration were activated in AAH relative to NL (Figure 3). Gene sets associated with enhanced cell cycle and proliferation as well as reduced apoptosis were modulated in LUAD relative to AAH or NL.

Differential expression programs between *BRAF*- and *KRAS*-mutant atypical adenomatous hyperplasia

We compared and contrasted gene expression among three groups of AAHs based on driver gene mutation status identified above (Figure 2): *BRAF* mutant, *KRAS* mutant and *BRAF/*

KRAS wild type (WT). Using a similar model as described above for global gene expression analysis but with a *P* < 0.01 threshold and a 1.5 fold-change cut-off, we identified 327 differentially modulated genes between the three different groups of AAHs (Figure 4; Table S7). Accordingly, these gene features indeed clustered the three groups separately based on the driver mutation status but with *BRAF*- and *KRAS*-mutant AAHs grouped closer together than with *BRAF*/*KRAS* WT AAHs (Figure 4; Table S7). Among the genes that were enriched in the *BRAF*-mutant AAHs is the cytokinesis promoting gene *KIF5C* and the cell proliferation promoting transcription factor (MYC Associated Factor X) *MAX*, typically associated with *MYC* oncoprotein [16] (Figure 4). On the other hand, *KRAS*-mutant AAHs displayed up-regulated expression of tumor necrosis factor receptor superfamily members 9 and 10B (*TNFRSF9* and *TNFRSF10B*), the NF-kB subunit *RELB* and the proliferation promoting ubiquitin ligase *UBE2C* (Figure 4). Of note, both *BRAF*-mutant and *KRAS*-mutant AAHs exhibited suppressed expression of the epithelial mesenchymal transition-promoting tyrosine kinase receptor *AXL* relative to *BRAF*/*KRAS* WT AAHs (Figure 4). These findings suggest shared and disparate expression programs among AAHs with activating mutations in the oncogenic GTPases *BRAF* and *KRAS*.

Profiles of immune function in the pathogenesis of atypical adenomatous hyperplasia

Accumulating evidence suggests a pivotal role for the host immune response in the evolution of cancer as well as dynamic interplay between emerging tumor cells and immune-based expression programs [17,18]. We sought to begin to understand contextual immune marker profiles in the development and progression of AAHs. Among these profiles we identified genes with known roles in immune signaling based on an annotated and *a priori* list from the nCounter PanCancer Immune Profiling Panel (nanoString Technologies). Using a similar random effects model described above, a significance threshold *P* < 0.001 and a 1.5 fold-change, we identified 131 markers of immune response that were differentially modulated among NLs, AAHs and LUADs (Figure 5; Table S8). Overall, the immune markers followed similar patterns or clusters of expression described above. This analysis revealed that *IL12A*, a cytokine most notably associated with an anti-tumor immune response [19], was decreased in AAHs and LUADs relative to NLs (Figure 5; Table S8). Conversely, the cytokines *CXCL13* and *CXCL14*, indicative of activated B-cell chemotaxis and signaling [20,21], were up-regulated in AAHs and LUADs (Figure 5). Moreover, we found aberrant immune marker expression occurring early in AAHs, relative to NLs (Figure 5). We found early and significantly decreased expression of prototypical markers of the anti-tumor immune response (e.g., *GZMB*) in AAHs relative to NLs (Figure 5). On the other hand, AAHs exhibited increased expression of the tumor-supportive chemokine receptor *CCR2* (Figure 5) [22]. Of note, we found that the major immune checkpoint cytotoxic T-lymphocyte-associated antigen 4 (*CTLA-4*) [23] was significantly up-regulated in LUADs relative to AAHs but not so in the premalignant lesions relative to NLs (Figure 5, Table S8) suggesting that aberrant immune checkpoint function by *CTLA-4* may be implicated in progression of AAH to LUAD. These findings accentuate the role of aberrant immune function and signaling early on in the development and progression of AAH.

DISCUSSION

There is a lack of understanding of the molecular aberrations leading to the initiation as well as the progression of atypical adenomatous hyperplasia (AAH), the only known precursor lesion to LUAD [3]. Here, we probe the mutation and gene expression landscapes of AAHs in comparison to normal tissues and early-stage LUADs from the same (matched) patients. We delineated subgroups of AAHs with mutually exclusive and distinct driver gene mutation status; namely *BRAF*-mutant (both non-smokers and ever-smokers), *KRAS*-mutant (ever-smokers only) and *KRAS*/*BRAF* WT AAHs. By agnostic transcriptome sequencing analysis, we also identified various patterns of expression profiles and pathways in the molecular pathogenesis of AAH. Further analysis underscored markers of immune function that are significantly differentially modulated, early on, in AAH (down-regulation of *GZMB*) relative to normal lung (NL) as well as those deregulated in LUADs relative to AAHs (e.g. *CTLA-4*). Our study highlights early recurrent driver mutations, expression profiles and markers of immune response in AAH that offer a window to understand the molecular pathogenesis of these premalignant lesions.

In this cohort, we found that the *BRAF* oncogene was the most commonly mutated gene in AAH (four patients with p.K601E and one with p.N581S, Table S3) followed by *KRAS* (predominantly in codon 12). No *BRAF* variants were found in the paired LUADs. Of note, among AAHs, mutations in *BRAF* and *KRAS* were mutually exclusive. Whereas *KRAS*-mutant AAHs were from ever-smokers, *BRAF* mutations in AAHs occurred in both non-smokers and ever-smokers. The *BRAF* p.K601E variant has been previously noted in preneoplastic melanocytic lesions and melanomas *in situ* as well as in thyroid adenomas [24–26], thus pointing to the probable role of *BRAF* in early stages of oncogenesis (i.e., development of preneoplasia such as AAH). The *BRAF* p.K601E mutation was also found in small proportions of cancers of the thyroid, colon and skin [27]. This may suggest that an enrichment for this hotspot driver mutation highlights a crucial mechanism for AAH and LUAD pathogenesis. Indeed, studies by the TCGA [12] and our group [28] showed relatively infrequent (~3%) *BRAF* mutations in LUADs. Yet, its absence in our sample set of LUADs, including in tumor specimens from patients with *BRAF*-mutant AAH is intriguing. It cannot be neglected that this may, in part, be due to our relatively modest set of samples (further discussed below).

An intriguing finding from our study is the pattern of *BRAF* and *EGFR* mutations in the paired LUADs. In four of the five *BRAF*-mutant AAHs, the paired LUADs exhibited driver *EGFR* mutations (e.g., p.L858R). Conversely, LUADs of *KRAS*-mutant AAHs displayed several other driver mutations (*TP53*, *KRAS*) that are typically associated with smoking [3]. Our study also underscores previously uncharacterized properties of these AAH *BRAF* mutations, namely mutual exclusivity with *KRAS* and correlation with smoking patterns. Based on our findings on mutual exclusivity of *BRAF* and *KRAS* in AAHs along with the disparate patterns of mutations in the paired LUADs, it is plausible to suggest that there are divergent pathways in pathogenesis of these preneoplastic lesions. A schematic of this paradigm is represented in Figure 6. A similar divergent model to malignancy has also been recently described in the evolution of different melanoma subtypes from their precursor lesions [26].

The multiregion sequencing study by Izumchenko *et al* [8] supports several of our findings. First, their study noted *BRAF* as the most commonly mutated gene in AAH, albeit in a smaller cohort of six patients. Further, they identified the *BRAF* p.K601E variant in a single AAH sample at a very low variant allele frequency (2%). Second, the mutual exclusivity of *BRAF* and *KRAS* mutations that we observed in AAHs can also be inferred from their data. Third, in both studies, *EGFR* and *KRAS* were the only driver genes sharing the exact mutation in both AAHs and LUADs. Fourth, among the 28 genes we found to be mutated in both AAHs and LUADs (Figure 2B), they observed three (*TP53, EGFR* and *KRAS*) to be mutated. Additionally, *APC, CDKN2A, CREBBP* and *NF1* which we observed as shared mutated genes were only reported in either AAHs or malignant lesions in their study. These disparities are not absolute as they may result from imperfect detection due to technological limitations (e.g. sequencing depths) and differences in study design. In comparison to their study, we surveyed a greater number of patients and genes as well as achieved a higher sequencing depth to detect potentially rare (within-sample) variants. Further, we identified recurrent *BRAF* mutations (most commonly p.K601E) that exhibited molecular (mutual exclusivity with *KRAS*) and clinicopathological (also found in non-smokers) features and that were present at a higher allele frequency (10 – 37%, Table S3) than observed previously. Nonetheless, given the complexity of AAHs and the malignant lesions leading to invasive LUAD, it is interesting, and indeed confirmatory, to identify similar gene mutation patterns across studies.

Complementing our DNA analysis is our agnostic transcriptome sequencing analysis that revealed differential gene expression programs that occur in different stages of LUAD pathogenesis – early in development of AAH from normal tissue, in LUADs, or in both lesion types. Gene set enrichment and pathway analysis pinpointed elevated immune cell trafficking and *WNT/*β-catenin signaling as well as an inhibition of the anti-tumor inflammatory response (Th1) and transforming growth factor beta 1 (*TGFB1*) signaling in development of AAH from normal lung. *WNT/*β-catenin signaling has been previously shown to be activated in progression of oral leukoplakia, a precancerous lesion of head and neck squamous cell carcinoma (HNSCC) [29]. Conversely, gene sets associated with increased cell cycle and proliferation, decreased apoptosis, and reduced function of pulmonary surfactant proteins (e.g., *SFTPs A1, A2, C*, and *D*) were enriched in profiles differentially expressed in LUAD relative to AAH. Several key signaling pathways (e.g. those mediated by *EGFR, MYC, CSF2*) were elevated/enriched in AAH and furthermore in LUAD, consistent with their role in tumor progression [3]. Increased *EGFR* was shown to promote cellular proliferation, inhibit apoptosis and drive development and progression of bronchial dysplasia [30]. Similarly, *MYC* overexpression has been previously reported in colorectal polyps with a level of expression proportional to the polyp size as well as dysplastic histology [31]. By transcriptome sequencing, we also identified differentially expressed profiles between AAHs with mutations in *BRAF* and *KRAS*. *BRAF*-mutant AAHs showed significant downregulation of adenosine deaminase (ADA), an enzyme involved in nucleotide metabolism, the deficit of which may lead to impaired DNA synthesis and repair [32]. In contrast, *SOS2*, an oncogene known to confer increased growth potential of tumor cells exhibiting an oncogenic *KRAS* [33], and ubiquitin conjugating enzyme E2C (*UBE2C*), previously shown to be elevated in lung cancer lesions, particularly in ever-

smokers [34], was enriched in the *KRAS*-mutant AAHs compared to *BRAF*-mutant AAHs. Taken together, these data point to early changes in the development and progression of AAH and that would thus comprise ideal targets for chemoprevention of LUAD.

Our findings on aberrant immune-regulated pathways from the agnostic gene expression analysis prompted us to more closely probe the modulation of known markers of immune function. Our immune marker profiling overall suggested an activation of pro-tumor immune pathways (i.e., Th2) and B-cell receptor signaling as well as an inhibition of anti-tumor immune response (e.g., Th1-derived *IFN-γ* signaling). Similar findings have been reported in previous studies of Barrett's esophageal tissues, a premalignant condition with a high risk of progression to esophageal adenocarcinomas [35]. *IL12A*, known for its proinflammatory anti-tumor response, along with anti-tumor immune chemokines (e.g. *CCL3, CCL4, TLR4*) and apoptosis-inducing proteases (*GZMB*) were decreased in AAH relative to normal lung. On the other hand, we found elevated expression of the *CCL2* chemokine receptor *CCR2* in AAH relative to normal lung. *CCR2* has been shown to enhance tumor growth, angiogenesis and tumor progression and was demonstrated to be over-expressed in several tumor tissues [22]. Recently, *CCL2/CCR2*-based immune prevention models were shown to attenuate tumor development and metastasis [22,36]. Of note, we identified an increasing expression in chemokines *CXCL13* and *CXCL14*, both known for their role in inflammatory processes and immune response [20], and *SPP1*, previously shown to be overexpressed in premalignant lesions of the oral epithelium as well as actinic keratosis (AK), the premalignant lesion to skin squamous cell carcinomas [37]. We also found that *CD27*, which in combination with its ligand *CD70* is known to generate a potent co-stimulatory signal, was increased in AAH relative to normal lung. Notably, our analysis pointed to significantly increased expression of the major immune checkpoint *CTLA-4* in LUAD relative to AAH [23]. We also observed disparate patterns of immune marker expression among AAHs with different recurrent mutated driver genes (*BRAF*-mutant, *KRAS*-mutant and *BRAF/KRAS* WT). We noted decreased expression of *TNFRSF9*, also called CD137, known to regulate the activation of T cells and a promising target for enhancing antitumor immune responses [38], in *BRAF*-mutant AAHs suggesting a relatively dampened immune response in this subgroup of AAH. Conversely, we found decreased expression of the receptor tyrosine kinase *AXL*, previously shown to promote pro-tumor immune responses [39], in AAHs with mutations in either *BRAF* or *KRAS*. These findings, albeit based on a limited cohort of AAHs analyzed by both DNA and RNA sequencing, point to differential aberrant immune signaling among AAH based on driver mutation status. Future studies surveying a larger number of AAH have the potential to corroborate these observations. Nonetheless, we posit here that aberrant immune signaling (e.g., attenuated anti-tumor immune response) is a common, perhaps critical, feature of AAH and LUAD development, as illustrated in Figure 6. Indeed, ongoing whole-exome sequencing studies have begun to shed light on neoantigen profiles in AAH [40], analogous to observations recently made for LUADs [41]. Also, further studies examining protein (e.g. by multiplexed immunohistochemistry) levels of markers of various immune cell infiltrates will shed more light on the role of the immune response in the pathogenesis of AAH. It is important to mention that immune-based therapy has come to the forefront of targeted therapeutic strategies for various malignancies including those of the lung [42]. For instance, monoclonal antibodies targeting genes such as

CD27 and *CTLA-4* are already being tested for treatment of various cancers including lung adenocarcinomas [43]. In this context and as alluded to previously [44,45], our findings suggest that targeting immune responses and signaling (e.g., immune checkpoint blockade) may be a viable strategy to prevent progression of preneoplasia such as AAH.

It is important to note that our study is not without limitations. While we comprehensively studied paired AAHs and LUADs, our cross sectional study design is not best positioned to thoroughly characterize the "progression" of AAH to LUAD. Naturally, the AAHs that already progressed to LUADs are no longer available for analysis. Our present report lends to the need for further longitudinal studies with a larger number of samples in which expression data and markers of the immune response can be aligned with time and space. Additionally, future longitudinal studies surveying AAHs in patients without overt lung malignancy may also help pinpoint drivers of AAH progression. Further, based on our findings on the absence of *BRAF* mutations in the LUADs studied in our cohort, one cannot rule out the possibility that the *BRAF*-mutant AAHs are benign and may not be the preneoplastic lesions that eventually progress to LUADs. Earlier studies have insinuated that *BRAF* mutations are important for initiation of premalignancy rather than their expansion or progression [8,46]. Lungs of mice genetically engineered to express a mutant form of *Braf* (p.V600E) were shown to develop hyperplasias that progressed to adenoma [47]. Of note, only after mutations in other genes (e.g. *Tp53*) did the *Braf* mutant lesions progress to LUAD [47]. Yet, the strong pairing of *BRAF*-mutant AAHs with *EGFR*-mutant LUADs is nonetheless an interesting observation that is worth investigating in future studies comprising a larger number of patients with both AAHs and LUADs. Further, that these patterns hold across lesions arising independently, although potentially from the same cell lineage, reflect the patient-specific nature of their development and highlight the potential for personalized prevention strategies. It is also worthwhile to mention that our cohort was mainly comprised of East Asian patients. Earlier studies have demonstrated that LUADs of East Asians exhibit disparate mutational spectra (e.g. more prevalent activating mutations in *EGFR*) relative to LUADs from Western (or Caucasian) patients [48,49]. It is reasonable to surmise that mutational differences in AAHs, across patients of different ethnicities, to roughly reflect those we observe in LUADs. In this context, our results and proposed paradigm could be more relevant to the East Asian population based on our cohort, and necessitates future work in larger populations comprising diverse ethnicities. Also, recent pathological classification guidelines for LUAD have underscored subgroups with pure lepidic growth (adenocarcinoma *in situ*; AIS) and those that exhibit predominant lepidic growth and with less than 5 mm invasion (minimally invasive adenocarcinoma; MIA) [3]. Earlier work in East Asian LUAD patients suggested that LUADs of the "terminal respiratory unit" (TRU) progress from AAH to AIS and then to invasive lesions [3]. It is plausible that AIS may have distinct profiles that suggest an intermediate stage [8,26] in the progression of AAH to LUAD. Our cohort largely comprised LUADs with very few AIS or MIA, too limited in size to further delineate profiles along this progression. Future studies are warranted to align with space mutational profiles, gene expression and markers of the immune response, particularly those shared between AAH and LUAD, and determine their contextual role in development of AIS and their progression to LUAD. Also, indeed, recent

and ongoing efforts have begun to distinguish potentially distinct profiles in the multi-step progression of AAH to AIS and subsequently to LUAD [3,8,40].

Findings from our study shed light on some of the earliest mutation events, expression changes as well as altered immune pathways in the pathogenesis of AAH, the only known precursor lesion to LUAD. The different mechanisms in the pathogenesis of AAH we explore here may further identify novel biomarkers and potentially offer immune-based intervention or other personalized prevention in patients with early-stage LUAD. This further accentuates the need for a greater depth and understanding of immunotherapeutic strategies early on, in potentially less hostile environments such as those typified by premalignant lung lesions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Herbst RS, Heymach JV, Lippman SM. Lung cancer. N Engl J Med. 2008; 359:1367–1380. [PubMed: 18815398]

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. CA Cancer J Clin. 2017; 67:7–30. [PubMed: 28055103]

3. Kadara H, Scheet P, Wistuba II, Spira AE. Early Events in the Molecular Pathogenesis of Lung Cancer. Cancer Prev Res. 2016; 9:518–527.

4. Mori M, Rao SK, Popper HH, Cagle PT, Fraire AE. Atypical Adenomatous Hyperplasia of the Lung: A Probable Forerunner in the Development of Adenocarcinoma of the Lung. Mod Pathol. 2001; 14:72–84. [PubMed: 11235908]

5. Sakamoto H, Shimizu J, Horio Y, Ueda R, Takahashi T, Mitsudomi T, et al. Disproportionate representation of KRAS gene mutation in atypical adenomatous hyperplasia, but even distribution of EGFR gene mutation from preinvasive to invasive adenocarcinomas. J Pathol. 2007; 212:287–294. [PubMed: 17534846]

6. Licchesi JDF, Westra WH, Hooker CM, Herman JG. Promoter hypermethylation of hallmark cancer genes in atypical adenomatous hyperplasia of the lung. Clin Cancer Res. 2008; 14:2570–2578. [PubMed: 18451218]

7. Selamat SA, Galler JS, Joshi AD, Fyfe MN, Campan M, Siegmund KD, et al. DNA methylation changes in atypical adenomatous hyperplasia, adenocarcinoma in situ, and lung adenocarcinoma. PLoS One. 2011; 6:e21443. [PubMed: 21731750]

8. Izumchenko E, Chang X, Brait M, Fertig E, Kagohara LT, Bedi A, et al. Targeted sequencing reveals clonal genetic changes in the progression of early lung neoplasms and paired circulating DNA. Nat Commun. 2015; 6:8258. [PubMed: 26374070]

9. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013; 31:213–219. [PubMed: 23396013]

10. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 2012; 22:568–576. [PubMed: 22300766]

11. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer Genome Landscapes. Science American Association for the Advancement of Science. 2013; 339:1546–1558.

12. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014; 511:543–550. [PubMed: 25079552]

13. Xing Y, Yu T, Wu YN, Roy M, Kim J, Lee C. An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. Nucleic Acids Res. 2006; 34:3150–3160. [PubMed: 16757580]

14. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47. [PubMed: 25605792]

15. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. J Thorac Oncol. 2015; 10:1243–1260. [PubMed: 26291008]

16. Dang CV. MYC on the path to cancer. Cell. 2012; 149:22–35. [PubMed: 22464321]

17. Gajewski TF, Schreiber H, Fu Y-X. Innate and adaptive immune cells in the tumor microenvironment. Nat Immunol. 2013; 14:1014–1022. [PubMed: 24048123]

18. Disis ML. Immune Regulation of Cancer. J Clin Orthod. 2010; 28:4531–4538.

19. Tugues S, Burkhard SH, Ohs I, Vrohlings M, Nussbaum K, vom Berg J, et al. New insights into IL-12-mediated tumor suppression. Cell Death Differ. 2014; 22:237–246. [PubMed: 25190142]

20. Lu J, Chatterjee M, Schmid H, Beck S, Gawaz M. CXCL14 as an emerging immune and inflammatory modulator. J Inflamm. 2016; 13:1.

21. Panse J, Friedrichs K, Marx A, Hildebrandt Y, Luetkens T, Barrels K, et al. Chemokine CXCL13 is overexpressed in the tumour tissue and in the peripheral blood of breast cancer patients. Br J Cancer. 2008; 99:930–938. [PubMed: 18781150]

22. Lim SY, Yuzhalin AE, Gordon-Weeks AN, Muschel RJ. Targeting the CCL2-CCR2 signaling axis in cancer metastasis. Oncotarget. 2016; 7:28697–28710. [PubMed: 26885690]

23. Grosso JF, Jure-Kunkel MN. CTLA-4 blockade in tumor models: an overview of preclinical and translational research. Cancer Immun. 2013; 13:5. [PubMed: 23390376]

24. Afkhami M, Karunamurthy A, Chiosea S, Nikiforova MN, Seethala R, Nikiforov YE, et al. Histopathologic and Clinical Characterization of Thyroid Tumors Carrying the BRAF(K601E) Mutation. Thyroid. 2016; 26:242–247. [PubMed: 26422023]

25. Macerola E, Torregrossa L, Ugolini C, Bakkar S, Vitti P, Fadda G, et al. BRAFK601E Mutation in a Follicular Thyroid Adenoma. Int J Surg Pathol. 2017:106689691668808.

26. Shain AH, Yeh I, Kovalyshyn I, Sriharan A, Talevich E, Gagnon A, et al. The Genetic Evolution of Melanoma from Precursor Lesions. N Engl J Med. 2015; 373:1926–1936. [PubMed: 26559571]

27. Zheng G, Tseng L-H, Chen G, Haley L, Illei P, Gocke CD, et al. Clinical detection and categorization of uncommon and concomitant mutations involving BRAF. BMC Cancer. 2015; 15:779. [PubMed: 26498038]

28. Kadara H, Choi M, Zhang J, Parra ER, Rodriguez-Canales J, Gaffney SG, et al. Whole-exome sequencing and immune profiling of early-stage lung adenocarcinoma with fully annotated clinical follow-up. Ann Oncol. 2017; doi: 10.1093/annonc/mdx062

29. Ishida K, Ito S, Wada N, Deguchi H, Hata T, Hosoda M, et al. Nuclear localization of beta-catenin involved in precancerous change in oral leukoplakia. Mol Cancer. 2007; 6:62. [PubMed: 17922924]

30. Merrick DT, Kittelson J, Winterhalder R, Kotantoulas G, Ingeberg S, Keith RL, et al. Analysis of c-ErbB1/Epidermal Growth Factor Receptor and c-ErbB2/HER-2 Expression in Bronchial Dysplasia: Evaluation of Potential Targets for Chemoprevention of Lung Cancer. Clin Cancer Res American Association for Cancer Research. 2006; 12:2281–2288.

31. Imaseki H, Hayashi H, Taira M, Ito Y, Tabata Y, Onoda S, et al. Expression of c-myc oncogene in colorectal polyps as a biological marker for monitoring malignant potential. Cancer. 1989; 64:704–709. [PubMed: 2743265]

32. Antonioli L, Blandizzi C, Pacher P, Haskó G. Immunity, inflammation and cancer: a leading role for adenosine. Nat Rev Cancer. 2013; 13:842–857. [PubMed: 24226193]

33. Jeng H-H, Taylor LJ, Bar-Sagi D. Sos-mediated cross-activation of wild-type Ras by oncogenic Ras is essential for tumorigenesis. Nat Commun. 2012; 3:1168. [PubMed: 23132018]

34. Kadara H, Lacroix L, Behrens C, Solis L, Gu X, Lee JJ, et al. Identification of gene signatures and molecular markers for human lung cancer prognosis using an in vitro lung carcinogenesis system. Cancer Prev Res. 2009; 2:702–711.

35. Kavanagh ME, Conroy MJ, Clarke NE, Gilmartin NT, O'Sullivan KE, Feighery R, et al. Impact of the inflammatory microenvironment on T-cell phenotype in the progression from reflux oesophagitis to Barrett oesophagus and oesophageal adenocarcinoma. Cancer Lett. 2016; 370:117–124. [PubMed: 26519754]

36. Jung H, Ertl L, Janson C, Schall T, Charo I. Abstract A107: Inhibition of CCR2 potentiates the checkpoint inhibitor immunotherapy in pancreatic cancer. Cancer Immunology Research. 2016; 4:A107–A107.

37. Chang P-L, Harkins L, Hsieh Y-H, Hicks P, Sappayatosok K, Yodsanga S, et al. Osteopontin Expression in Normal Skin and Non-melanoma Skin Tumors. J Histochem Cytochem Histochemical Society. 2008; 56:57.

38. Yonezawa A, Dutt S, Chester C, Kim J, Kohrt HE. Boosting Cancer Immunotherapy with Anti-CD137 Antibody Therapy. Clin Cancer Res. 2015; 21:3113–3120. [PubMed: 25908780]

39. Gay CM, Balaji K, Byers LA. Giving AXL the axe: targeting AXL in human malignancy. Br J Cancer. 2017; 116:415–423. [PubMed: 28072762]

40. Krysan, K., Tran, LM., Grimes, BS., Walser, TC., Wallace, WD., Dubinett, SM. Evaluation of progression associated neoepitopes and immune contexture in pulmonary premalignancy [abstract]. Proceedings of the 108th Annual Meeting of the American Association for Cancer Research; 2017 Apr 1–5; Washington, DC. Philadelphia (PA): AACR; 2017. Abstract nr 1016

41. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015; 348:124–128. [PubMed: 25765070]

42. Knutson KL, Disis ML. Tumor antigen-specific T helper cells in cancer immunity and immunotherapy. Cancer Immunol Immunother. 2005; 54:721–728. [PubMed: 16010587]

43. Cully M. Combinations with checkpoint inhibitors at wavefront of cancer immunotherapy. Nat Rev Drug Discov. 2015; 14:374–375. [PubMed: 26027531]

44. Spira A, Disis ML, Schiller JT, Vilar E, Rebbeck TR, Bejar R, et al. Leveraging premalignant biology for immune-based cancer prevention. Proc Natl Acad Sci U S A. 2016; 113:10750–10758. [PubMed: 27638202]

45. Young MRI. Redirecting the focus of cancer immunotherapy to premalignant conditions. Cancer Lett. 2017; 391:83–88. [PubMed: 28130162]

46. McCaskill-Stevens W, Pearson DC, Kramer BS, Ford LG, Lippman SM. Identifying and Creating the Next Generation of Community-Based Cancer Prevention Studies: Summary of a National Cancer Institute Think Tank. Cancer Prev Res. 2017; 10:99–107.

47. Dankort D, Filenova E, Collado M, Serrano M, Jones K, McMahon M. A new mouse model to explore the initiation, progression, and therapy of BRAFV600E-induced lung tumors. Genes Dev. 2007; 21:379–384. [PubMed: 17299132]

48. Li S, Choi Y-L, Gong Z, Liu X, Lira M, Kan Z, et al. Comprehensive Characterization of Oncogenic Drivers in Asian Lung Adenocarcinoma. J Thorac Oncol. 2016; 11:2129–2140. [PubMed: 27615396]

49. Krishnan VG, Ebert PJ, Ting JC, Lim E, Wong S-S, Teo ASM, et al. Whole-genome sequencing of asian lung cancers: second-hand smoke unlikely to be responsible for higher incidence of lung cancer among Asian never-smokers. Cancer Res. 2014; 74:6071–6081. [PubMed: 25189529]
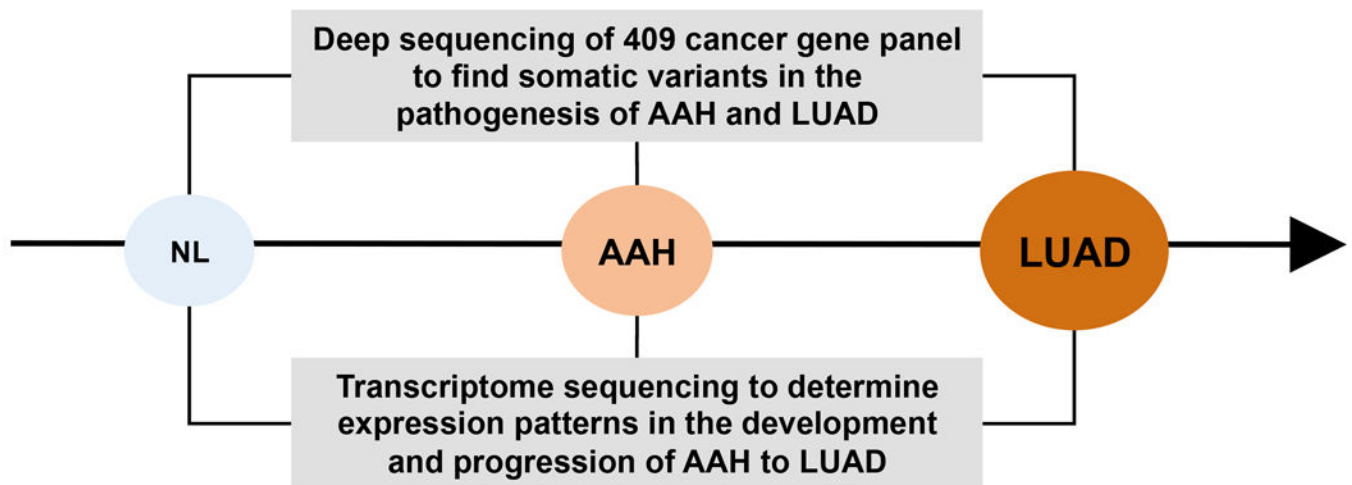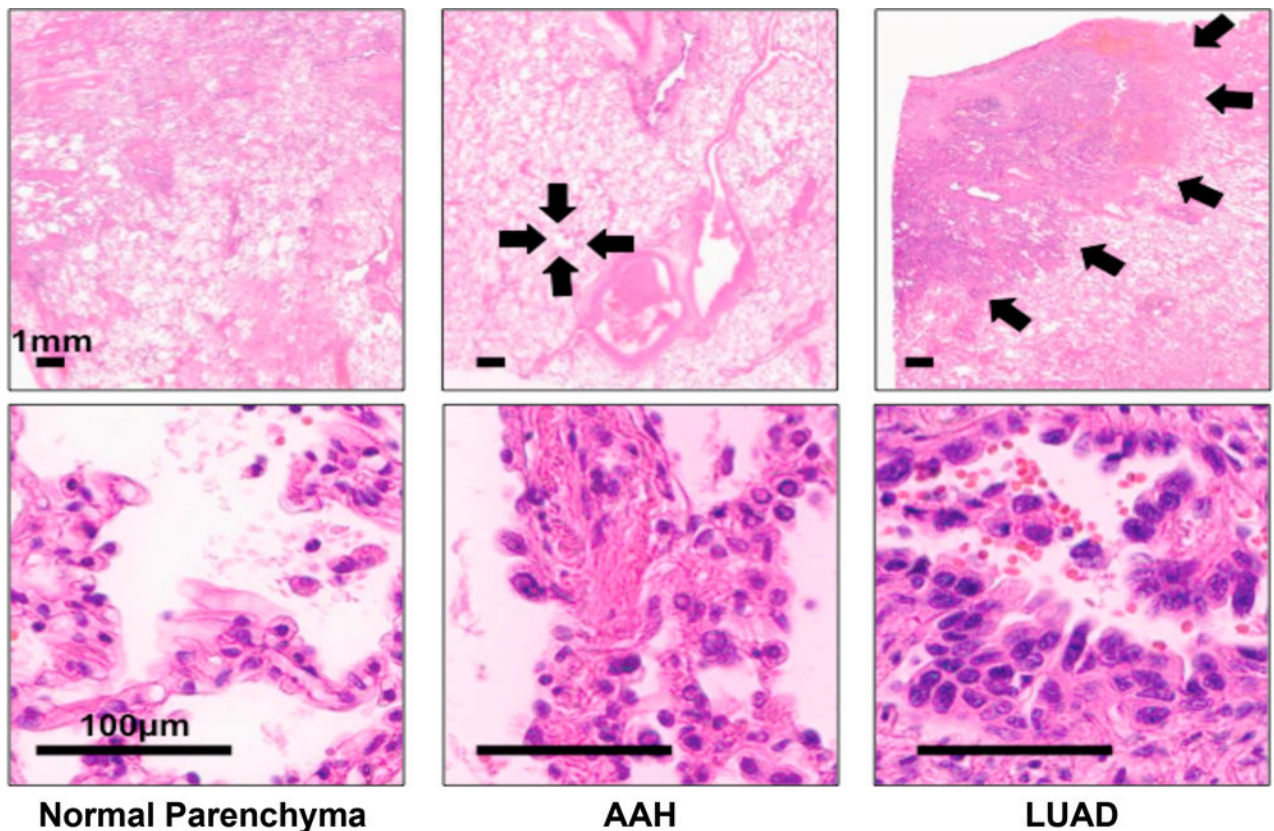
Figure 1. Study design to understand the development and progression of adenomatous atypical hyperplasia

Diagnosis and histopathological determination of specimens following H&E staining was performed to determine and classify normal lung tissues (NL), atypical adenomatous hyperplasia (AAH) and lung adenocarcinoma (LUAD). A two pronged approach was used to study the pathogenesis of AAH. Deep targeted sequencing of 409 cancer-associated genes was performed to identify somatic point mutations and transcriptome sequencing was carried out to study expression profiles.
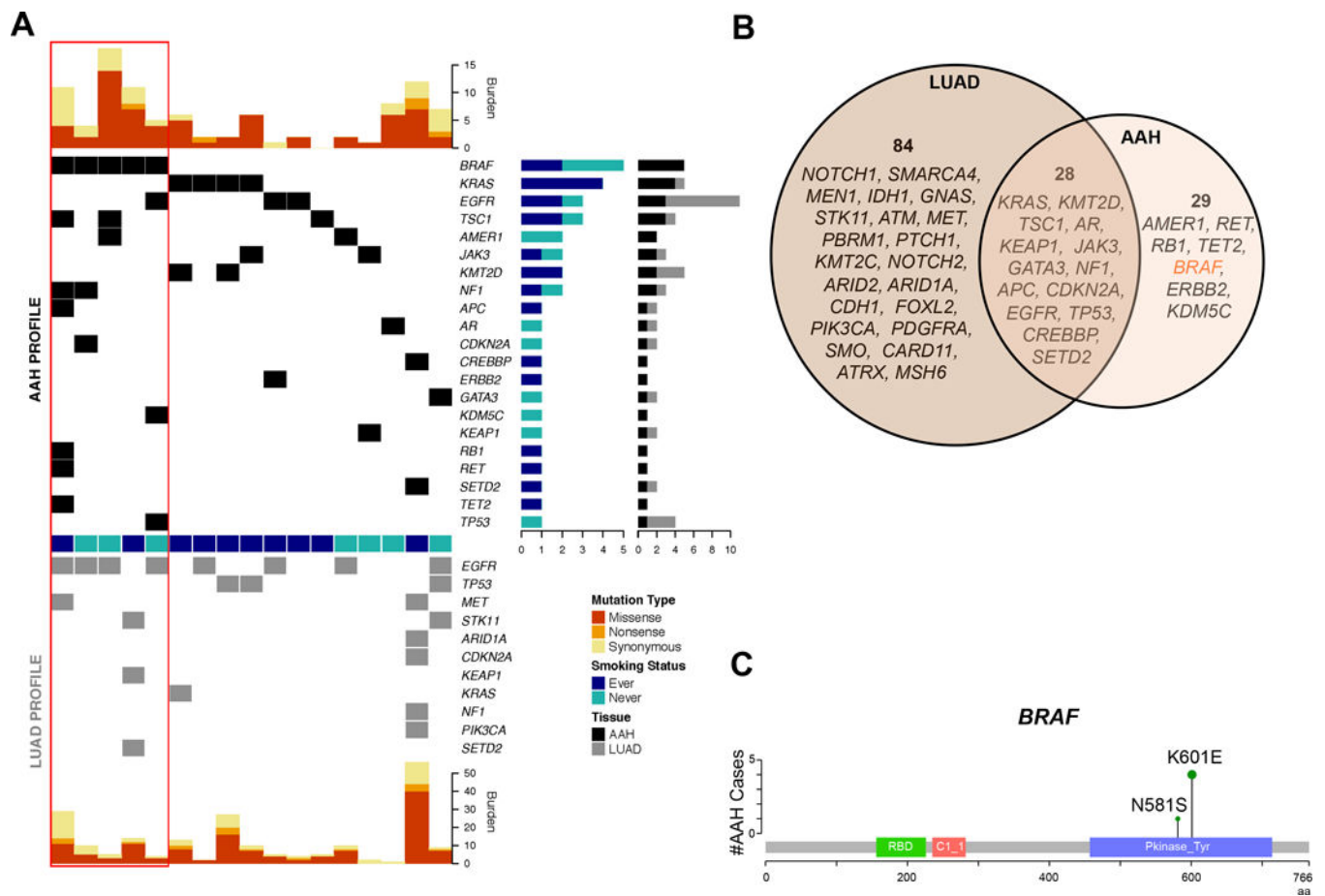
Figure 2. Somatic mutation profiles in atypical adenomatous hyperplasia

Deep targeted sequencing of a cancer gene panel ($n = 409$) and identification of somatic nonsynonymous mutations in AAHs and LUADs was performed as described in the Methods section. (A) We examined, in greater detail, mutations in previously established lung cancer drivers from the TCGA [12] as well as other known cancer-associated genes [11]. AAH specimens ($n = 17$) that exhibited a mutation in either driver gene set were plotted. The paired LUADs were also plotted depicting mutations in genes previously established by the TCGA to be significantly mutated in LUAD [12]. Shown within the red panel is the enrichment of *EGFR* mutations in LUAD (80%) paired to *BRAF*-mutant AAH. (B) A tissue level analysis of mutations in AAH and LUAD specimens was performed to identify mutated genes, from the same set of driver genes surveyed in panel A, that were common or disparate between AAH and LUAD. (C) Lollipop plot for mutations (p.K601E; $n = 4$ and p.N581S; $n = 1$) in the *BRAF* gene and their prevalence in AAH specimens.
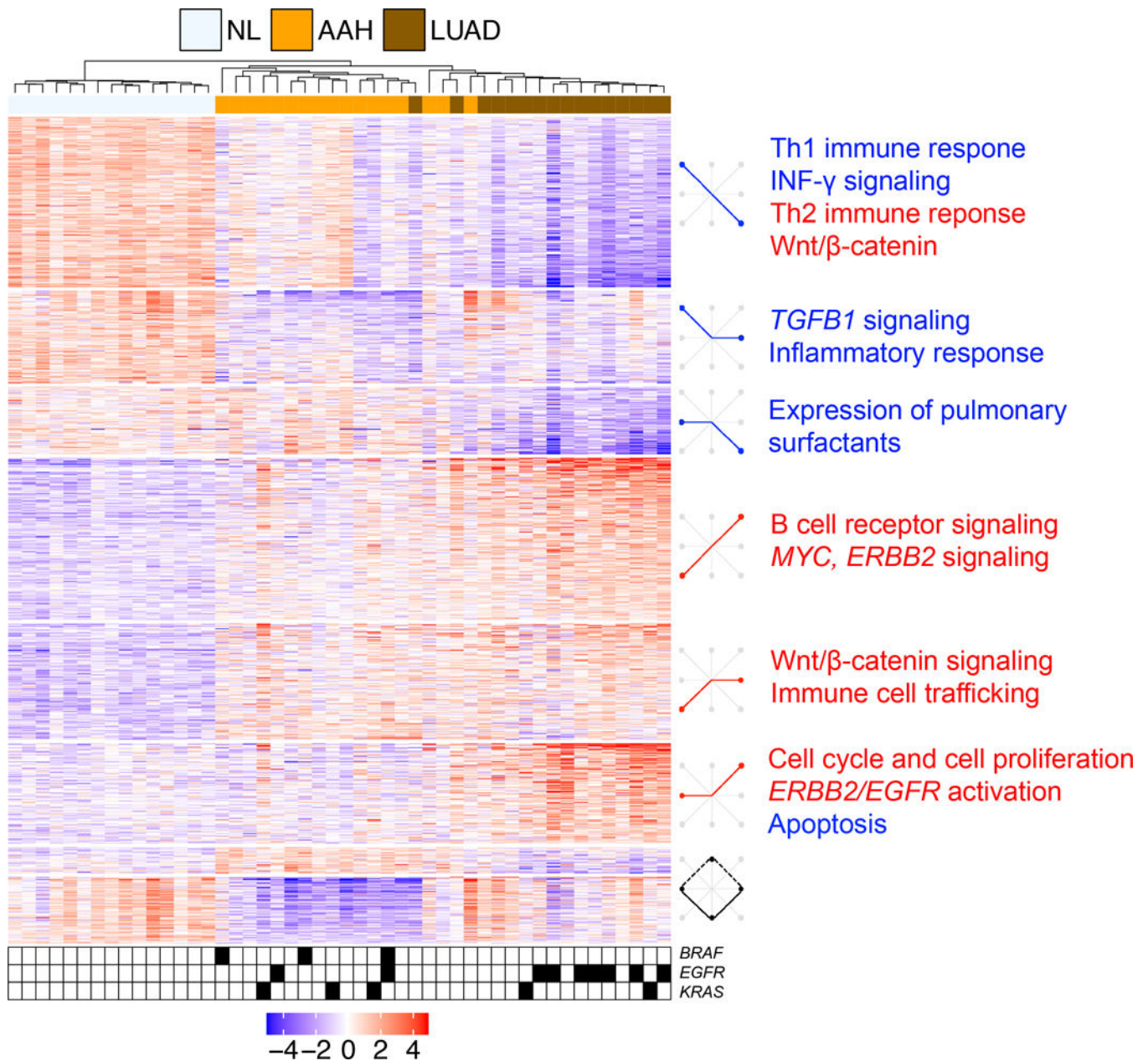
Figure 3. Expression profiles differentially modulated in development of atypical adenomatous hyperplasia and lung adenocarcinoma

Transcriptome sequencing was performed as described in the Methods section. Genes (n = 1008) differentially expressed between the three tissues (AAH vs NL, LUAD vs NL or LUAD vs AAH) were determined using ANOVA (P < 0.001, 2-fold change) and analyzed by hierarchical clustering (red, up-regulated relative to median sample; blue, down-regulated relative to median sample). Genes were grouped into eight different patterns based on two one-sided t-tests for NL to AAH and AAH to LUAD comparisons. Patterns of differential expression in each gene cluster are schematically depicted on the right. Pathways and gene set enrichment analysis was performed using Ingenuity Pathways Analysis. Pathways deregulated in each cluster of genes are depicted in red (activation) and blue (inhibition)
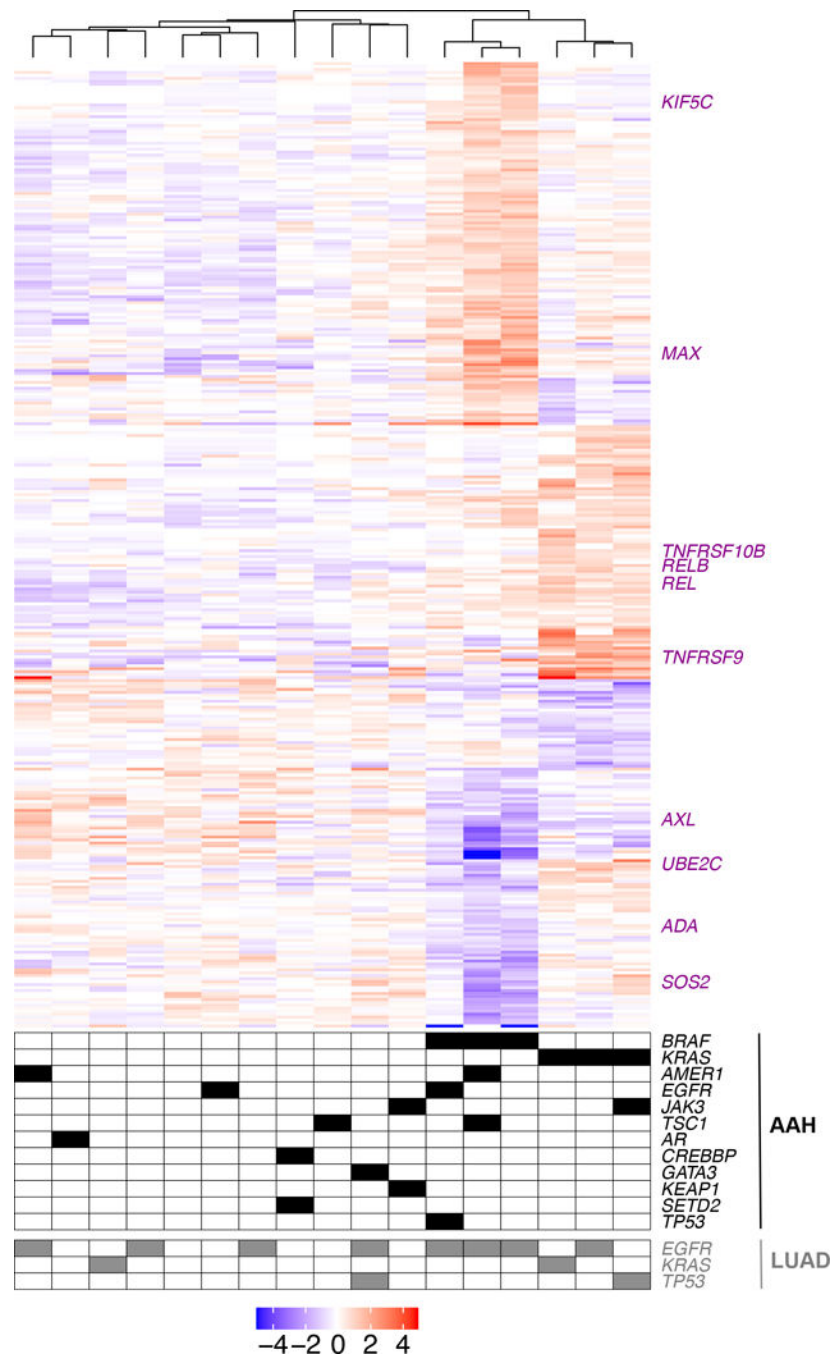
alongside the heatmap. Mutations status of *EGFR, KRAS* and *BRAF* for AAH and LUAD specimens is depicted below.

Figure 4. Differential gene expression based on driver mutation status in atypical adenomatous hyperplasia

AAHs were subgrouped based on *BRAF* and *KRAS* mutation status: *BRAF*-mutant, *KRAS*-mutant and *BRAF/KRAS* wild type. Genes ($n = 327$) differentially expressed between the three AAH subgroups were identified using ANOVA ($P < 0.01$, 1.5 fold-change) and analyzed by hierarchical clustering (red, up-regulated relative to median sample; blue, down-regulated relative to median sample).
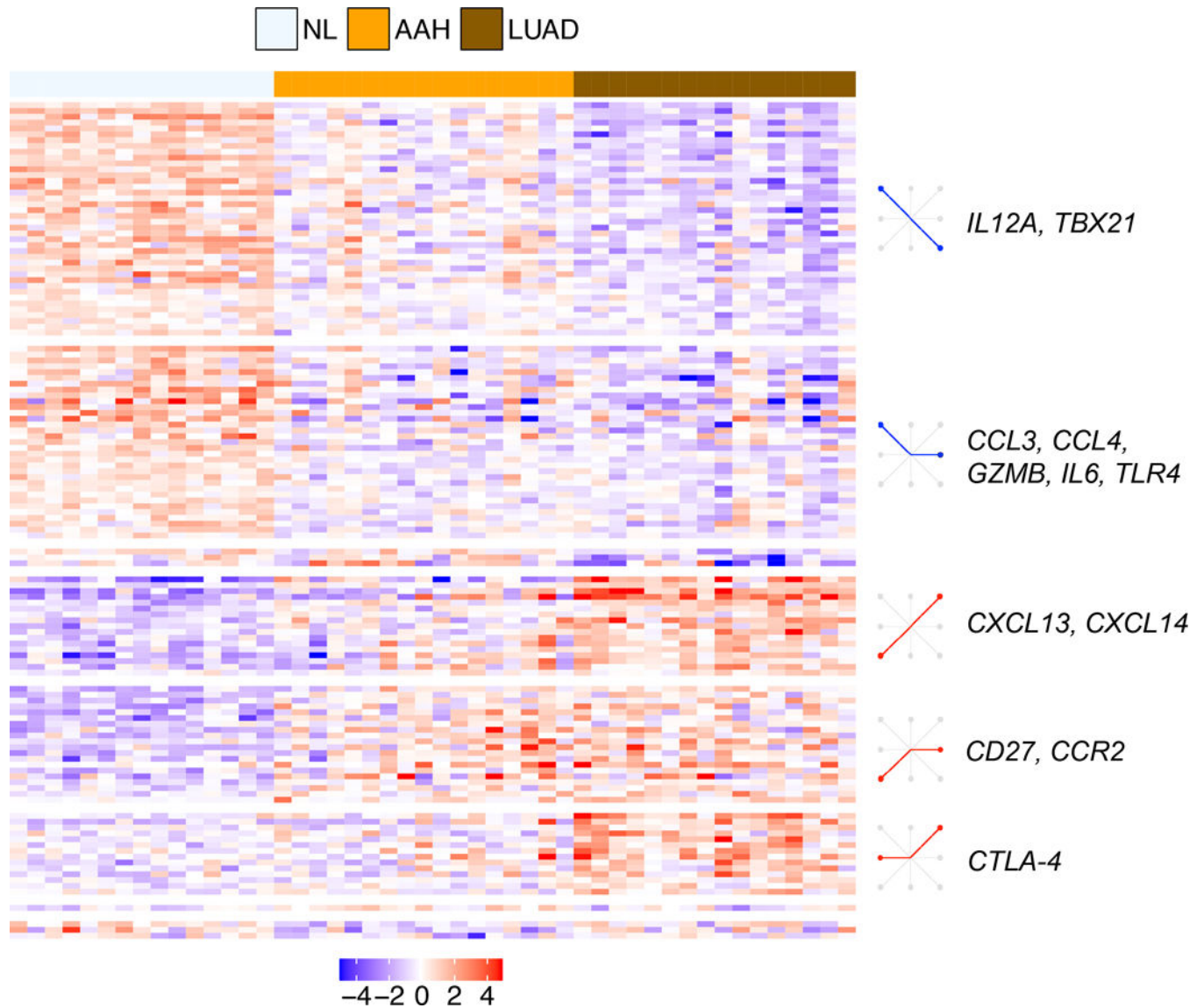
Figure 5. Deregulation of immune signaling in the molecular pathogenesis of atypical adenomatous hyperplasia

Expression profiles for an *a priori* list ($n = 730$) of immune markers from the nCounter PanCancer Immune Profiling Panel (nanoString technologies) was compiled (see Materials and Methods section) and studied to identify differentially expressed immune genes ($n = 131$; ANOVA; $P < 0.001$ and 1.5 fold-change). The genes were divided into different clusters based on patterns of differential expression between NL, AAH and LUAD derived from two one-sided *t*-tests (AAH vs NL and LUAD vs AAH). Patterns of differential expression in each gene cluster and select immune markers are schematically depicted on the right. present in major clusters are also depicted on the right.
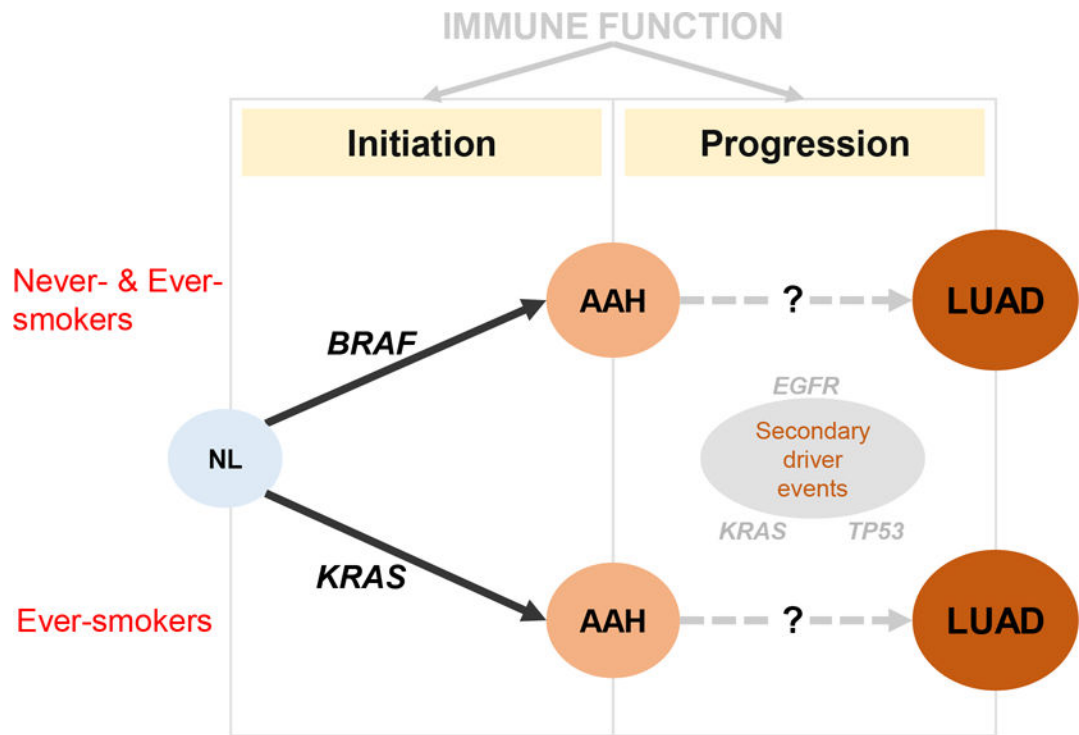
Figure 6. Proposed models for the pathogenesis of atypical adenomatous hyperplasia
Two potential divergent modes in the pathogenesis of these preneoplastic lesions are proposed based on the mutual exclusivity of mutations and disparate expression profiles. A subgroup of AAHs, occurring in both non-smokers and ever-smokers, are initiated by *BRAF* and tend to be associated with development of LUADs with driver mutations in the *EGFR* oncogene (not excluding the possibility that the *EGFR*-mutant LUADs may have arisen from different AAHs). Mechanisms involved in the potential progression of *BRAF*-mutant AAH to LUAD (e.g. *EGFR*-mutant tumors) warrant further studies. Another subset of AAHs are driven by *KRAS*, occur predominantly in ever-smokers and lead to LUADs with mutations in other driver genes besides *KRAS* (e.g., *TP53*). Transcriptome sequencing analysis pointed to aberrant immune signaling (e.g., up-regulated *CTLA-4*) in the pathogenesis of AAH. Further analysis (e.g. of a larger cohort of AAH) may help underscore profiles, immune markers and pathways unique to each molecular group of AAHs thus paving the way for new strategies (e.g., immune-based) for (chemo)prevention and early intervention.

Table 1

Cases studied with atypical adenomatous hyperplasia.

Case	Age	Gender	Tobacco history	Stage
1	70	Male	Ever	IA
2	21	Female	Ever	IB
3	67	Male	Ever	IA
4	46	Female	Ever	IA
5	79	Female	Never	IA
6	40	Male	Ever	IA
7	72	Male	Never	IA
8	48	Female	Never	IA
9	51	Female	Never	IB
10	81	Female	Never	IA
11	67	Male	Ever	IA
12	63	Female	Never	IA
13	79	Female	Ever	IA
14	54	Female	Never	IA
15	62	Male	Ever	IA
16	64	Female	Never	IA
17	67	Male	Ever	IA
18	57	Female	Never	IA
19	71	Male	Ever	IB
20	63	Male	Ever	IIIA
21	74	Female	Never	IA
22	60	Male	Ever	IA