

## AMERICAN THORACIC SOCIETY DOCUMENTS

Consensus Lung Cell Reference: The Collaborative Cross-Consortium and Country Lung Cell Nomenclature Project (C3LCN). An Official American Thoracic Society Workshop Report

Gloria S. Pryhuber, Bruce Aronow, Martin Banchero, Kevin Burns, Wellington V. Cardoso, Jeremy Clair, Gail Deutsch, Thu Elizabeth Duong, Joshua Fortriede, Gerard H. Koppelman, Jonathan A. Kropski, Raphael Kfuri-Rubens, Sara Lin, Malte D. Luecken, Susan Majka, Kerstin Meyer, Ravi S. Misra, Terren K. Niethamer, Anna Karina Perl, Jayaraj Rajagopal, Eniko Sajti, Nathan Salomonis, Richard H. Scheuermann, Xin Sun, Purushothama Rao Tata, Alexandra-Chloe Villani, Matthew S. Walters, Kathryn A. Wikenheiser-Brokamp, Laure-Emmanuelle Zaragosi, Martijn C. Nawijn; on behalf of the American Thoracic Society Assembly on Respiratory Cell and Molecular Biology

ORCID IDs: 0000-0002-9185-3994 (G.S.P); 0000-0001-5109-6514 (B.A.); 0009-0003-8398-4614 (M.B.); 0009-0003-6587-4809 (K.B.); 0000-0002-8868-9716 (W.V.C); 0000-0003-0478-8520 (G.C.); 0000-0002-0571-0285 (G.D.); 0000-0001-7122-4448 (T.E.D.); 0000-0002-3293-5463 (J.F.); 0000-0001-8567-3252 (G.H.K.); 0000-0002-8923-1344 (J.A.K); 0000-0003-3256-9640 (R.K-R.); 0000-0002-7559-1963 (S.L.); 0000-0001-7464-7921(M.D.L.); 0000-0002-1005-1552 (S.M.); 0000-0001-5906-1498 (K.M.); 0000-0002-1352-4378 (R.S.M.); 0000-0002-0914-994X (T.K.N.); 0000-0002-8445-4565 (A.K.P.); 0000-0002-4122-177X (J.R.); 0000-0002-0531-7715 (E.S.); 0000-0001-9689-2469 (N.S.); 0000-0003-1355-892X (R.H.S.); 0000-0001-8387-4966 (X.S.); 0000-0003-4837-0337 (P.R.T.); 0000-0001-7461-0408 (A-C.V.); 0000-0003-0010-8444 (M.S.W.); 0000-0002-2932-3095 (K.A.W-B.); 0000-0001-6747-7928 (L-E. Z.); 0000-0003-3372-6521 M.C.N.)

### Affiliations

Gloria S. Pryhuber, M.D.<sup>1</sup>  
 Martijn Nawijn, Ph.D.<sup>2</sup>  
 Denise Al Alam, Ph.D.<sup>3</sup>  
 Bruce Aronow, Ph.D.<sup>4</sup>  
 Martin Banchero, M.S.<sup>2</sup>  
 Nicholas Banovich, Ph.D.<sup>5</sup>  
 Pascal Barbry, Ph.D.<sup>6</sup>  
 Maria Basil, M.D., Ph.D.<sup>7</sup>  
 Kevin Burns, Ph.D.<sup>4</sup>  
 Janette Burgess, Ph.D.<sup>2</sup>  
 Wellington V. Cardoso, M.D.<sup>8</sup>  
 Jeremy Clair, Ph.D.<sup>9</sup>  
 Rachel Clifford, Ph.D.<sup>10</sup>  
 Soula Danopoulos, Ph.D.<sup>3</sup>  
 Gail Deutsch, M.D.<sup>11</sup>  
 Thu Elizabeth Duong, M.D.<sup>12</sup>  
 Joshua Fortriede, Ph.D.<sup>4</sup>  
 David Frank, M.D., Ph.D.<sup>7</sup>  
 Reinoud Gosens, Ph.D.<sup>2</sup>  
 Mingxia Gu, Ph.D.<sup>13</sup>  
 Minzhe Guo, Ph.D.<sup>4</sup>  
 James Hagood, M.D.<sup>14</sup>  
 Raphael Kfuri-Rubens, M.D., M.Phil.<sup>15, 16</sup>

Gerard H. Koppelman, M.D., Ph.D.<sup>2</sup>  
 Stephanie Krick, M.D., Ph.D.<sup>17</sup>  
 Jonathan A. Kropski, M.D.<sup>18</sup>  
 Qing Sara Lin, Ph.D.<sup>19</sup>  
 Clare Lloyd, Ph.D.<sup>20</sup>  
 Malte D. Luecken, Ph.D.<sup>15</sup>  
 Susan Majka, Ph.D.<sup>21</sup>  
 Kerstin Meyer, Ph.D.<sup>22</sup>  
 Alexander Misharin, M.D., Ph.D.<sup>23</sup>  
 Ravi S. Misra, Ph.D.<sup>1</sup>  
 Ana Mora, M.D.<sup>24</sup>  
 Enid Neptune, M.D.<sup>25</sup>  
 Terren K. Niethamer, Ph.D.<sup>26</sup>  
 Kenichi Okuda, M.D., Ph.D.<sup>14</sup>  
 David Osumi-Sutherland, Ph.D.<sup>27</sup>  
 Anna Karina Perl, Ph.D.<sup>4</sup>  
 Joseph D. Planer, M.D.<sup>7</sup>  
 Aleix Puig-Barbé, Ph.D.<sup>27</sup>  
 Ellen M. Quardokus, M.S.<sup>28</sup>  
 Rod Rahimi, M.D., Ph.D.<sup>29</sup>  
 Jayaraj Rajagopal, M.D.<sup>29</sup>  
 Scott Randell, Ph.D.<sup>14</sup>  
 Elizabeth Redente, Ph.D.<sup>21</sup>  
 Eniko Sajti, M.D., Ph.D.<sup>12</sup>  
 Nathan Salomonis, Ph.D.<sup>4</sup>  
 Christos Samakovlis, Ph.D.<sup>30</sup>  
 Richard Scheuermann, Ph.D.<sup>31</sup>  
 Janine Schniering, Ph.D.<sup>32</sup>  
 Xin Sun, Ph.D.<sup>12</sup>  
 Purushothama Tata, Ph.D.<sup>33</sup>  
 Alexandra-Chloé Villani, Ph.D.<sup>29</sup>  
 Matthew S. Walters, Ph.D.<sup>34</sup>  
 Kathryn A. Wikenheiser-Brokamp, M.D., Ph.D.<sup>4</sup>  
 Yan Xu, Ph.D.<sup>4</sup>  
 Laure-Emmanuelle Zaragosi, Ph.D.<sup>6</sup>

<sup>1</sup> University of Rochester Medical Center, Rochester, New York, USA

<sup>2</sup> University of Groningen / University Medical Center Groningen, Groningen, Netherlands

<sup>3</sup> Harbor-UCLA Medical Center / Lundquist Institute for Biomedical Innovation, Torrance, California, USA

<sup>4</sup> Cincinnati Children's Hospital Medical Center / University of Cincinnati College of Medicine, Cincinnati, Ohio, USA

<sup>5</sup> Translational Genomics Research Institute (TGen), Phoenix, Arizona, USA

<sup>6</sup> CNRS / Institute of Molecular & Cellular Pharmacology, Université Côte d'Azur, Valbonne, France

<sup>7</sup> Perelman School of Medicine, University of Pennsylvania / Penn-CHOP Lung Biology Institute, Philadelphia, Pennsylvania, USA

<sup>8</sup> Columbia University Irving Medical Center, New York, New York, USA

<sup>9</sup> Pacific Northwest National Laboratory, Richland, Washington, USA

<sup>10</sup> University of Nottingham, Nottingham, United Kingdom

<sup>11</sup> Seattle Children's Hospital / University of Washington, Seattle, Washington, USA

<sup>12</sup> University of California San Diego, La Jolla, California, USA

<sup>13</sup> University of California Los Angeles, Los Angeles, California, USA

<sup>14</sup> Marsico Lung Institute, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

<sup>15</sup> Helmholtz Munich, Munich / Neuherberg, Germany

- <sup>16</sup> Technical University of Munich, Munich, Germany
- <sup>17</sup> University of Alabama at Birmingham, Birmingham, Alabama, USA
- <sup>18</sup> Vanderbilt University Medical Center, Nashville, Tennessee, USA
- <sup>19</sup> National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, USA
- <sup>20</sup> Imperial College London / Francis Crick Institute, London, United Kingdom
- <sup>21</sup> National Jewish Health, Denver, Colorado, USA
- <sup>22</sup> GSK / University of Cambridge, Cambridge, United Kingdom
- <sup>23</sup> Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA
- <sup>24</sup> Ohio State University Wexner Medical Center, Columbus, Ohio, USA
- <sup>25</sup> Johns Hopkins University, Baltimore, Maryland, USA
- <sup>26</sup> Yale School of Medicine, New Haven, Connecticut, USA
- <sup>27</sup> European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom
- <sup>28</sup> Indiana University, Bloomington, Indiana, USA
- <sup>29</sup> Massachusetts General Hospital / Harvard Medical School / Broad Institute, Boston, Massachusetts, USA
- <sup>30</sup> Stockholm University, Stockholm, Sweden
- <sup>31</sup> National Library of Medicine, National Institutes of Health, Bethesda, Maryland, USA
- <sup>32</sup> University of Bern / Helmholtz Munich, Bern, Switzerland
- <sup>33</sup> Duke University, Durham, North Carolina, USA
- <sup>34</sup> University of Central Oklahoma, Edmond, Oklahoma, USA

This official Workshop Report of the American Thoracic Society was approved May 2026.

You may print 1 copy of this document at no charge.

This document was funded by the American Thoracic Society.

Artificial Intelligence Disclaimer: AI (Chat-GPT Version 1.2026.160) was used to augment a literature search for top impactful lung cell biology references, primary and reviews, since 1900. The search results were then trimmed and curated by authors for relevance to lung cell identification.

## ABSTRACT

Rapid advances in single-cell technologies now allow measurement of thousands of transcripts and other molecular features of individual cells offering unprecedented insight into lung biology in homeostasis and in disease. The accelerated generation of multimodal data has, however, been accompanied by the reporting of putatively “novel” cell types described without consensus regarding their ontogeny, identity, function, or defining markers. To fully realize the value of the technological advances and to enable rigorous comparison across studies, respiratory research will benefit from standardized, quantitative, and biologically grounded cell classifications and nomenclature. Achieving the transformative potential of the multimodal data will depend on common, machine- and human-readable nomenclature, structured and expandable dictionary and atlas resources, and clear methodological standards that ensure consistency as technologies evolve.

The American Thoracic Society (ATS) recognized the importance of promoting a common nomenclature to enhance equitable access and utility of the vast amounts of multimodal data generated by the lung research community. The Collaborative Cross-Consortium and Country Lung Cell Nomenclature Project (C3LCN) was adopted as an ATS Assembly Project in 2024. This is the consensus report outlining the goals and framework of the Project to foster coordinated progressive lung cell research to include: 1) providing best practices for analysis, publication and reporting of lung single-cell transcriptomic datasets; 2) establishing a contemporary lexicon for healthy adult human cells of the lower respiratory tract with structured, persistent, and resolvable identifiers; 3) defining a scalable taxonomy to organize a common lung cell nomenclature; 4) offering tools to support collaboration, knowledge dissemination, and translational advances rooted in modern lung biology augmenting, not replacing, pre-genomic biological knowledge; and 5) describing an infrastructure capable of incorporating new ontological refinements as higher-resolution, multimodal single-cell and spatial datasets emerge, cellular heterogeneity is better defined, and disease-associated abnormal cell types and reactive cell states are increasingly recognized and mechanistically interrogated. Together, this coordinated effort aims to provide the foundation necessary for a robust, harmonized, and expandable nomenclature for lung science.

**Keywords:** RNAseq, transcriptomics, taxonomy, ontology, lung cell annotation

## CONTENTS

Overview

*Scientific Knowledge on the Subject*

*What This Study Adds to the Field*

Key Concepts and Conclusions

Introduction

Methods

Current Lung Cell Atlas Efforts

Lung Cell Database Resources

Standards for Single-cell Data Publication

The Cross-Consortium Lung Cell Nomenclature Workshop

Current and Evolving Consensus Lung Cell Nomenclature

Summary

## OVERVIEW:

### *Scientific Knowledge on the Subject*

Rapid advances in single-cell and spatial profiling have generated unprecedented insight into lung biology but also a proliferation of inconsistently defined lung cell “types” and “states.” The absence of a harmonized nomenclature limits reproducibility, complicates dataset integration, and slows translation of cellular discoveries into mechanistic understanding and clinical impact.

### *What This Study Adds to the Field*

This ATS consensus report establishes a coordinated framework for lung cell nomenclature, including structured and machine-readable annotations, best practices for reporting single-cell datasets, a scalable taxonomy for lung cell identities, and shared tools and infrastructure for community-wide use. This effort provides the foundation for a unified lung cell nomenclature that will enhance synergy across studies and accelerate collaborative discovery in pulmonary research.

## Key Concepts and Conclusions

- Rapid advances in single-cell and spatial technologies now enable measurement of thousands of molecular features per cell, generating unprecedented insights into lung biology at homeostasis and under perturbation conditions.
- The proliferation of novel cell-type labels across studies, often without shared definitions or lineage context, has created barriers to comparison, reproducibility, and translation. A harmonized nomenclature and consensus for its maintenance will improve the quality and speed of respiratory research.
- A shared, machine-readable, and human-interpretable ontology for lung structures and cell types is essential for FAIR data use (Findable, Accessible, Interoperable and Reusable), optimizing data value, translation, and open science.
- Best-practice recommendations are provided for annotating and publishing lung single-cell and spatial transcriptomic datasets, including approaches to standardized metadata, nomenclature conventions, and requirements for transparency and reproducibility.
- A flexible taxonomy framework is provided that prioritizes cell identities expected in healthy adult human lung and that will accommodate novel cell identities found in homeostatic conditions as well as dynamic, disease-associated states, enabling alignment across experimental platforms and biological contexts.
- This ATS-sponsored consensus nomenclature effort demonstrates what can be accomplished by bringing together experts from 18 institutions across five countries, spanning basic biology, computational science, physiology, and clinical pulmonary medicine, including but not exclusive to consortia members, to reach scientific consensus.
- Tools and infrastructure are introduced to support and encourage continued community-driven curation, allowing the lung cell reference framework to expand and adapt as new high-dimensional datasets and biological insights emerge.
- Overarching goals and framework are identified to accelerate discovery and translation in pulmonary science and medicine, by enabling researchers to use a common nomenclature when defining, comparing, and interrogating lung cell populations.

## INTRODUCTION

Recent advances in omics technologies have enabled high-resolution profiling of hundreds to thousands of gene transcripts, epigenomic marks, proteins, lipids, and metabolites at single-cell resolution. Increasingly, these methods also support *in situ* spatial analysis, revealing the remarkable heterogeneity of cell identities, including cell types, sub-types, and states, as well as their organization into functional multicellular units within tissues (functional tissue units - FTUs) (See Glossary, Table 1). Large-scale efforts to map cells and create tissue, organ, and system atlases, together with the development of advanced model systems, have uncovered a far greater diversity of molecularly defined cell identities in healthy organs than previously recognized. Harmonization across studies, needed to confirm expanded cellular diversity, is however complicated by variation in tissue composition between atlases and by the dependence of cell-type and marker-gene identification on the cellular populations represented in each dataset. Consistent cross-study annotation is, as a result, challenging. Substantial phenotypic variation exists, even among healthy humans, influenced by factors such as age, sex, ancestry, and cumulative exposures. Cross-species differences in cell-type specific transcriptomic features further complicate cell type identification and translation of insights from model organisms to human health and disease. These sources of variability challenge classical concepts of broad canonical cell types as stable entities and underscore that cell identity must be defined across multiple levels and dimensions. Though interpretation is complex, the expanding molecular knowledge provides a unique foundation for identifying normal developmental stage alterations occurring across the lifespan, as well as in response to perturbations, with significant implications for diagnosis and therapeutic innovation regarding disease.

Defining transient and stable cell identities based on molecular features alone is complicated by many factors including discordance across omics layers (e.g. RNA vs protein), limitations inherent to molecular assays, choice of computational methods, and differences in research focus and expertise across annotating scientists. Reaching agreement even at the level of taxonomy, for example whether to define hierarchical cell identities by lineage, function, location, morphology, class, type, or state is challenging. As a result, inconsistencies in cellular identity annotations have occurred across single-cell datasets and references that impede understanding in high resolution lung biology. (1, 2) The value in reaching consensus on lung cell

nomenclature and cell-type specific signatures derived from single-cell RNAseq studies of healthy lung samples is increasingly well recognized. Applying standards to data quality assessment and analysis, along with the use of accurate yet flexible cell nomenclature, ontology, and taxonomy will advance community-wide efforts to critically appraise large scale, multi-source biological data and foster reproducible science. In addition, a shared ontology-supported nomenclature improves clarity and consistency in human communication but also enhances the performance of machine learning and other artificial intelligence (AI) approaches that depend on structured and harmonized vocabularies to achieve optimal results. (3) This ATS Assembly Workshop report provides recommendations for terminology to be used in this effort (Table 1), an initiating taxonomy (Table 2), and a framework for iterative and collaborative nomenclature development.

## METHODS

To address the issue of establishing a standardized, cross-consortium, common terminology compendium to support the lung community, The Lung Cell Summit was initiated and hosted at the University of California San Diego by the LungMAP consortium, supported by the NHLBI, in May 2022. Members of the LungMAP consortium were joined by national and international HCA and independent investigators with stature in the lung biology field. Four task force working groups, focused on epithelial, endothelial, stromal/mesenchymal, and immune cells, were formed to be inclusive of junior and senior pulmonary biologists, pathologists, computational biologists and informaticians. Their charge was to take a data driven approach to cataloging lung cell types, to work collectively, and to base decisions on primary data and published literature. The LungMAP CellCards resource was expanded into a Lung Cell Nomenclature database and populated with lung cell types and RNA transcript markers as reported by major resources published between 2020 and 2023 (1, 2, 5, 8, 12, 20). Additional references, indicated as published or self-reported, were added by working groups as appropriate to provide both mouse and human data to support cell type identifications. Cell Ontology IDs and descriptions were linked into the database for live synchronization. A pre-workshop teleconference “Charting the Lung: Annotations and Ontologies Towards a Consensus Lung Cell Atlas,” hosted by the ATS Section on Genetics and Genomics, provided an occasion to engage the lung research community and to identify challenges, opportunities, and approaches to integrate molecular data from otherwise independent projects. With this preparation, the ATS Workshop was held in May 2024, providing

an opportunity for in-person (n=34) and virtual (n=6) discussions among working group task force leads to identify topics for focus and for an impressive presentation of the HCA Cell Annotation Platform and plans for cell identity jamborees. The robust discussions that followed the working group summaries contributed to this Report to complete the Workshop project phase. Additional Lung Cell Nomenclature work was then achieved in an NIH sponsored 2-day virtual workshop that attracted international participants using a Miro board, an online whiteboard system used by teams to visualize ideas, brainstorm, and work on projects collaboratively. (79) This LCN collaboration whiteboard was preconstructed with cell superclass playgrounds designed to promote discussion. Outlined goals for the workshop were to specify positive and negative consensus markers (genes or proteins) for existing cell types, to specify descriptions for existing cell types, to specify the “categorical” parent cell type for subtypes, to propose experiments to gather missing data, to propose new cell subtypes of parent canonical cell types, and to specify data-driven marker gene combinations for both existing and novel cell types identified in single cell experiments. Participants ranged from students and junior investigators to well-known leaders in the field, many of whom openly shared data, experience, and questions. Ontology curators and bio-informaticians were specifically invited and provided guidance, infrastructure, and connections to reference NCBO ontologies and the LCN and HCA databases. Working group co-chairs led the discussions and second day follow-up. Participants were encouraged to interact verbally and to leave sticky notes on the board with recommendations, questions, and insights. These recorded comments were extracted and included in the product of this Report and the on-going work of the LCN. To augment communication, the LCN collaboration whiteboard remains active and accessible.

In addition, in recognition of the large body of “pre-genomics” foundational lung research community work on cell type identification and characterization upon which the LCN is built, AI (Chat-GPT Version 1.2026.160) was used to augment a literature search for top impactful lung cell biology references, primary research and reviews, since 1900, ranked within 30-50 year epochs and ensuring coverage of Level 1 cell types. Approximately 300 references were returned and available upon request. For the Workshop report, the search results were trimmed and curated by authors for relevance to lung cell identification. The search requests used were: “Provide a historian-level canonical lung cell biology bibliography, ranking the top 25 most important lung cell biology papers in each epoch 1900–1950; 1950-1980, 1981-2010, 2011-2026.” Then

added: “Find the impactful lung cell biology review articles from 1980-2026 looking for at least two for each of epithelial cells, endothelial cells, immune cells, and mesenchymal cells. The articles should be specific for lung.” The result represents an anthology of key articles in pulmonary biology and demonstrates the expertise brought to the LCN through these publications and in person by LCN subject matter experts.

This ATS Assembly Project Report was drafted from recordings and notes taken at each meeting, shared, and revised with the writing group members to reach its submitted form. The consensus effort is ongoing. Working groups continue to meet and are working on cell specific manuscripts. The anticipated integrated HCA Lung Bionetwork atlas (ILCA) is available on CAP for consensus annotation. Updates on the consensus annotations and the cell identity features achieved will be incorporated into the LCN, to remain available via the LungMAP portal and through periodic publications.

## RESULTS

### Current Lung Cell Atlas Efforts

Several synergistic lung cell atlas efforts are currently running in parallel including efforts of the National Heart, Lung, and Blood Institute’s Molecular Atlas of Lung Development Program ([LungMAP](#)), the [Lung Biological Network](#) of the Human Cell Atlas (HCA) consortium, the Human BioMolecular Atlas Program ([HuBMAP-Lung](#)), and the lung components of the Senescence Network ([SenNet](#)), as well as many initiatives from individual investigators (Table 3, Supp Table 1). In 2014, a need for cataloging of nomenclature in the investigational Lung community was recognized by the LungMAP consortium. Progress was made with the publication of LungMAP “Anatomic Ontologies for Human and Mouse Lung Maturation” (4) and “[Cell Ontology for Human Lung Maturation](#)”. These first-in-kind Lung Cell Ontologies helped to create the framework for the “Census of Lung: Cell Cards from LungMAP” publication (5) and the [CellCards](#) resource at LungMAP.net. Similarly, the HCA consortium, in 2016, initiated an international, collaborative effort aiming to define all human cell types by distinctive patterns of gene expression, physiological states, developmental trajectories, and spatial relationships in tissue. (6) The HCA Lung Bionetwork and the LungMAP consortium then published landmark resources based on integration of multiple, originally independent, transcriptomic datasets (7-18) with the intent to provide reference cell atlas resources for healthy human lung and for models in use for

studies in disease and across species. “An Integrated Cell Atlas of the Lung in Health and Disease” reported the integrated Human Lung Cell Atlas (HLCA v.1) (2) that combined 14 human respiratory system datasets into an atlas representing 107 healthy individuals and 584 thousand cells with an extended atlas containing an additional 32 datasets and a total of 2.2 million cells from healthy and diseased lung. The LungMAP “Guided Construction of Single Cell Reference for Human and Mouse Lung,” or “CellRef,” informed identification of highly curated cross-species lung cell types, 48 in human and 40 in mouse, starting with “seed” populations that best represented the CellCards dictionary of cell identities. (1, 5) In parallel, the HuBMAP consortium developed “Anatomical Structures, Cell Types and Biomarkers” references for Lung, Trachea and Main Bronchi that report cell types relying on manually curated gene and protein markers from both HLCA.v1, CellRef and literature searches for established biology, placing each cell in context of anatomical, radiological, and pathological standards. (19-23) While there is considerable agreement between these large consortium efforts, inconsistencies, some minor and others more consequential, are noted in cell type names, cell subtype resolution level, recognition of cell subtypes defined by spatial location and declared cell specific marker genes and proteins. These observations when aligning the references to each other, as well as to publications of independent investigators, suggested the benefit of collaborations to harmonize a lung cell dictionary.

In 2022, the LungMAP consortium held a Lung Cell Summit to address the issue of establishing a standardized, cross-consortium, multi-modal common nomenclature compendium. Four task force working groups, focused on epithelial, endothelial, stromal/mesenchymal and immune cells, were formed to be inclusive of junior and senior scientists and to include pulmonary biologists, clinicians, pathologists, computational biologists, informaticians, as well as curators for the Uber Anatomy Ontology (Uberon) (24) and Cell Ontology (CL). (25) Their charge was to take a data driven approach to cataloging lung cell types, to work collectively, and to base decisions on primary data and published literature. To foster the project, the Lung Cell Summit project created a Lung Cell Nomenclature database hosted on the LungMAP website as a CellCards resource. This database provided a framework for the ATS Assembly Project.

In parallel, the HCA Lung Bionetwork is developing “An Updated Integrated Healthy and COPD Lung Cell Atlas reference (ILCA)”, an updated HLCA resource. The goals of the ILCA include the integration of HLCA

v.1 and CellRef datasets to support consensus annotations, improved granularity of cell-type annotation per lineage, an increased focus on immune cells, improved representation of underrepresented ancestries, and an intent to demonstrate clinically relevant association of molecular cell identities and profiles to clinical pulmonary function parameters and chronic obstructive pulmonary disease (COPD) phenotypes. As a result of this ATS Assembly Project, the Cross-Consortium and Country Lung Cell Nomenclature (C3LCN) Project, ILCA cell annotations will be harmonized with those of other Resources (Table 3) utilizing the LungMAP Lung Cell Nomenclature (LCN) database, Cell Ontology, and the Cell Annotation Platform (CAP).

### Lung Cell Database Resources

In the post-genomics era, multiple sites now store and make accessible molecular lung cell data including LungMAP, HuBMAP, Human Reference Atlas, HCA, CELLxGENE data portals, and others (Table 3). The LungMAP LCN database and the CAP are two resources gathering curated cell identity information. The Lung Cell Summit project created the LCN database and data gathering tool to encourage and provide a framework for the collection of species-specific (mouse and human) entries of cell names, synonyms, Cell Ontology (CL) IDs and descriptions, typological hierarchy, anatomical restrictions with Uberon IDs, cell phenotypes including function, morphology, location, biomarkers (transcriptomic with HGNC gene IDs, protein imaging with antibody RRIDs, and cytometric data), and disease phenotypes (with Mondo Disease Ontology, or similar ontological IDs) (Figures 1 and 2). The use of CL as a resource for single-cell omics translation has been recently favorably reviewed. (26) References, with DOI or PMID, to support data entries are expected as are measures of confidence in the knowledge reported. Cells that may be transient or unique to a few individuals and not yet validated can be entered for consideration as a provisional cell and included in the Provisional Cell Ontology. The LCN data tool is built in REDCap, a user-friendly interactive system, flexible in design and data entry and highly malleable to create structured reports. (27) Persistent IDs are auto filled and updated directly into the LCN database from National Center for Biomedical Ontology (NCBO) Bio Portal ontologies, including from the CL, Uberon, Mondo ontologies. Entries and revisions to the LCN occur through the task force working groups (Supplemental Table 1) as described for the ATS Project. All elements of the LCN are readily exported for use in education, coordination amongst investigators and in machine learning

(Figures 1 and 2). In the process of using this data gathering and consensus building tool, knowledge gaps are identified. Given the importance of murine transgenic and developmental models, the current LCN database is built to mirror human and mouse cells. The project as begun is meant to be useful for spatial as well as single cell data annotation, filling a need to reflect not only hierarchical origin/lineage clustering and similarities of cellular transcriptomic programming. Spatial data interpretation requires a method of discussing single cells in the context of their location, niche, neighborhood. The LCN was expanded and organized in anticipation of this need, to define cells by anatomical structure, cellular behavior, function, and other characteristics modified by location, as well as transcript clustering. This does add a level of complexity to the nomenclature. Some super classes, like epithelial, are amenable to structure specific definitions while others, like mesenchymal lineage cells may better be distinguished by behavior, e.g. contractile or connective, as well as location, while immune cells are best defined by their immunological function. The open database structure is also amenable to expansion to variations in organization. A cell lineage-based hierarchy, resulting in a different dendrogram organization, would better serve developmental research, while cells identified in lung injury and repair may be accommodated in a hybrid organization. Currently, information related to disease and lineage are captured in the LCN database and displayed in linkage with the dendrogram (Figure 2).

The CAP is a web-based workspace for scientists to collaboratively annotate cell types and metadata. Annotation of the integrated ILCA single cell data reference will be facilitated by combining expert insights via the CAP infrastructure. CAP allows users to define and manage cell names in a private space before publication to facilitate creating a standardized, hierarchical system for cell annotation across different datasets and to foster consensus within biological networks. It is designed as a community-driven platform to provide automated annotation with cell type prediction models, as well as to explore and store subject matter expert perceptions on annotation encouraging converging on consensus nomenclature and marker characteristics and to identify knowledge gaps. The LCN Cell Cards and the CAP efforts have synergistic strengths to extend molecular characterization modalities and, in large part due to this ATS Assembly project, will work together to build consensus and resources to meet the needs of the Lung research community. The developing [LCN database](#) and [CAP](#) resources are currently available to the public.

## Standards for Single-cell Data Publication

Before embarking on development of data-based expanded nomenclature, it is important to understand the caveats of the technologies that are broadening our catalog of identified lung cell types and subtypes. The quality of nomenclature depends on the accuracy and reliability of the data generated and published. Current state-of-the-art efforts toward cell identification, as presented in this document, rely heavily on decisions made in reporting cellular transcriptomic assays. As Supplemental Material, we provide a brief primer on methods of generating and analyzing single-cell and single nucleus RNA sequencing (sc- or snRNAseq) resulting in “transcriptomic datasets” intended to capture gene expression profiles at single-cell level to resolve cellular heterogeneity.

## The Cross-Consortium Lung Cell Nomenclature Workshop

The ATS Assembly Project and related international efforts aim to strengthen cell identity annotations and promote clear, harmonized communication among investigators worldwide. It is increasingly clear that the impact of molecular mapping is magnified when cumulative data are organized within a shared, consensus-based reference framework. Cell type annotation assigns biological identities to clusters of cells that share a coherent molecular profile, while taxonomy and nomenclature provide the standardized terminology and hierarchical relationships needed to describe those identities consistently. Together, they support findable, accessible, interoperable, and reusable (FAIR) (28) data practices, enhance biological interpretation, improve cross-dataset comparability, and enable regulatory-grade reproducibility. The ATS 2024 C3LCN Workshop, sought to promote the development of a consensus, jointly determined process, common nomenclature, and resources to be used in annotation of lung cell identities, preferably considering not only single cell / nuclei RNA sequencing but, ultimately, occurrence of proteins and other metabolites resolved to location and anatomical structures. Focus topics identified and discussed at the Workshop by Task Force leaders (Supplemental Table 1) are summarized here.

Broadly relevant, substantial limitations were first noted that temper recently published cell classification. Most proposed cell types are defined solely by scRNAseq transcriptional signatures, without confirmation at

the protein level or functional validation in vivo. Differences in tissue digestion protocols, cell class enrichment strategies, and the number of captured cells across data sets introduce technical biases that may obscure or inflate distinctions between cell subtypes. Moreover, transcriptomic similarity does not necessarily imply lineage relatedness or shared physiological roles. Thus, rigorous cross-platform validation, including proteomics, flow cytometry, in situ hybridization, and functional assays, is still needed to resolve which cell identities represent stable lineages, dynamic states, environmentally induced phenotypes, or technical artifacts. Also acknowledged was the occurrence of transient cell states that may be novel or may resemble novel or other distinct cell types. It is important to “name the thing” for further exploration, especially considering the potential for these states to play a role in disease and repair, yet the potential for transient state or masquerade is to be remembered and determined.

Consensus supported that refining the classification of lung cells will deepen mechanistic understanding of lung biology and its role in maintaining pulmonary health. A more precise and validated lung cell atlas has the potential to reveal novel biomarkers, uncover therapeutic targets for injury and repair defined at a single cell level, and provide insight into the microenvironmental cues that govern immune regulation, gas exchange, and tissue remodeling. As the field continues to integrate multi-omics approaches with advanced modeling and experimental validation, we are poised to uncover a more comprehensive and dynamic view of lung cell identity, one that captures not only the cell types that constitute stable lung structures, but also transitions, states, and adaptive responses that enable this complex network to sustain life.

Task force specific topics include:

**EPITHELIAL:** The epithelial working group concluded that substantial consensus exists for several classical human lung epithelial cell types grounded in anatomy and long-standing functional studies, including basal cells, multiciliated cells (following correction from “ciliated”), alveolar epithelial type I (AT1) and type II (AT2) cells, as well as rare populations such as pulmonary neuroendocrine cells, tuft cells, and the more recently described ionocytes. In contrast, challenges remain for epithelial populations whose identities appear more heterogeneous or state-dependent in single-cell datasets. These include basal cell–related populations such as suprabasal, parabasal and hillock-associated cells, for which lineage relationships and defining criteria remain unclear, as well as deuterosomal cells, whose status as a transient precursor to multiciliated

cells versus a distinct cell type could use further resolution. Discussion also pointed out that consensus nomenclature for club and airway secretory cells is unclear, with ongoing debate as to whether club and goblet cells represent distinct types or context-dependent states along a shared secretory continuum, especially in inflammatory or injury settings. Updated nomenclature is needed for submucosal gland epithelial cells, highlighting the need to incorporate morphology, anatomical location, and molecular features to distinguish mucous and serous populations. Multiple groups have also described overlapping secretory cell populations in terminal and respiratory bronchioles, underscoring the need for harmonization across studies. Also in recent years, several investigators identified novel epithelial cell types including fibrosis-associated “aberrant basaloid” cells and subsets in human respiratory bronchioles that are quite similar between studies, yet it remains unclear if these cells and subsets represent identical cell types, and thus ontological synonyms, or condition-specific cell states. (29-31) The LCN, as it expands into disease, can be used to address these challenges, to pursue a coordinated, iterative effort to align annotations across datasets, track synonyms to accommodate differences in publications, define consensus marker genes, and update relevant cell ontologies as needed, with the goal of establishing a practical, interoperable framework that can evolve alongside emerging single-cell and spatial multimodal technologies.

**ENDOTHELIAL:** The endothelial group recognized several areas for refinement in lung endothelial cell identities, recent advances in single-cell transcriptomics having transformed understanding of endothelial cell (EnC) diversity in the adult, non-diseased lung across humans, mice, and related species. That the lung contains three vascular beds exposed to quite different fluids (e.g., oxygenated vs. deoxygenated vs. without red blood cells). The pulmonary circulation vasculature, systemic circulation vasculature, and lymphatic vasculature are clearly called out for probable distinction in endothelial cells. A foundational consensus is recognized that now organizes lung EnCs into several major subgroups: macrovascular endothelium, including pulmonary arterial, pulmonary venous, and systemic arterial and venous types, microvascular endothelium comprising capillary type 1 (Cap1) and capillary type 2 (Cap2, aerocyte) cells, and a distinct lymphatic endothelial population. This emerging framework provides a crucial scaffold for exploring endothelial specialization yet is thought to represent only the first layer of a far more intricate cellular landscape. A growing body of evidence suggests that each of these broad EnC categories contains

substantial and functionally meaningful heterogeneity. Endothelial cells exhibit remarkable positional specialization along the vascular tree, shaped by local hemodynamic forces, metabolic demands, tissue microenvironments, and organ-specific interactions with epithelial, stromal, and immune cells. These contextual cues give rise to EnC subpopulations that differ not only in transcriptional signatures but also in barrier properties, vasoactive responses, antigen presentation potential, and regenerative capacity. As such, refining EnC classification is not merely a taxonomic exercise, it is central to understanding how vascular niches regulate inflammation, maintain lung homeostasis, and orchestrate repair after injury.

A major future challenge is to identify and validate functional EnC subpopulations that may exist transiently or at very low abundance. Rare reparative, progenitor-like, or injury-associated endothelial states may exert disproportionate influence on processes such as alveolar regeneration, fibrosis, vascular remodeling, and angiogenesis. Capturing these states sufficiently to add to lung cell nomenclature will require methodological innovation, better preservation of fragile cell types during tissue processing, strategic enrichment approaches, lineage-tracing models, and integrative analyses that combine transcriptomic, epigenomic, proteomic, and spatial data. Complementing scRNAseq with spatial transcriptomics and advanced imaging will be particularly important for understanding how EnC states map to specific vascular niches and how these niches change across health, aging, and disease. In the meantime, to ease communication, predominant states can be added to LCN as provisional cells.

**IMMUNE:** The immune cell working group recognized that standardizing lung immune-cell nomenclature poses unique challenges distinct from those for epithelial, endothelial, or mesenchymal compartments. Immune cells are highly heterogenous, phenotypically dynamic, and mobile cell populations. Immune cells originate outside of the lung, either in the fetal yolk sac or liver as is the case for alveolar macrophages, or from the bone marrow as for post-natal monocyte derived macrophages. It is essential then, but challenging, to distinguish resident, recently recruited, and circulating immune cells. These categories can share overlapping transcriptional signatures but differ fundamentally in origin and function. Pronounced regional, matrix, and cellular variations across the airways, parenchyma, vasculature, and draining lymph nodes further complicate classification yet trafficking, activation and suppression of immune cells are highly dependent on cellular neighborhoods. Immune cells exhibit exceptional activation-driven plasticity via common pathways,

meaning that transient states can closely resemble what are considered distinct cell types, more so than appears true for most structural lung cells. Variation in human exposures and subsequent immune cell reactions may blur cell identities in comparing tissue samples of individuals. Immune cells are also particularly susceptible to processing artifacts. Many key immune subsets such as neutrophils, eosinophils, basophils, and mast cells are sensitive to tissue dissociation such that demonstrating their residence in tissue, especially by RNAseq, can be unreliable. Stress responses induced during tissue dissociation can also obscure the homeostatic cell identity. Classical immune cell definition additionally requires reconciling transcriptomic annotations with typically antigen-based, marker and gating lineage. Single-cell RNA datasets confirm canonical markers previously identified by protein may be insufficient transcriptomic markers such that unsupervised clustering by transcriptomics can be misleading. (32, 33) Among immune cell markers, CD4, for example, a strong marker for T helper lymphocytes by protein, is low or frequently even discordant when measured by RNAseq. (34, 35) Finally, spatial microenvironments including mucosal surfaces, glandular and vascular compartments, bronchus associated lymphoid tissue (BALT), and other lung immune cell niches, strongly and uniquely shape immune-cell identity and function, highlighting the need for nomenclature that captures both lineage, location, and context of cellular neighborhoods. (36) Given the heterogeneity, neighborhood-dependent states, and relatively rare but key immune cells found in the lung, (e.g., ILCs, MAIT cells, specialized dendritic cells, memory T cells), domain experts continue to work on identification and nomenclature of lung resident immune cell subtypes. Modifications in approved LCN items, currently relatively conservative, will evolve with new data. Renewed efforts to standardize immunology nomenclature, faced with ad hoc broadening of immune cell sub-setting due to advancing technologies and experimental models, are underway and will be considered in LCN updates to maintain consistency within and across fields. (37-41)

**MESENCHYMAL / STROMAL:** The terms *mesenchymal*, *stromal*, and *fibroblast* are frequently used interchangeably in the lung literature but are more accurately understood as hierarchical concepts. Consistent with Cell Ontology and legacy definitions, *Mesenchymal cells* are non-polarized precursor cells associated with the cell matrix and capable of migration. *Stromal cells* are connective tissue cells found in the loose connective tissue, derived from the mesenchymal lineage, and providing a microenvironment that supports

organ parenchymal cells. *Fibroblasts* represent a subset of stromal cells characterized by connective tissue functions and substantial tissue- and state-specific heterogeneity. Because of their pronounced spatial heterogeneity and capacity to adopt multiple functional states, fibroblasts were viewed during the workshop as a particularly stringent and informative test case for applying nomenclature principles. Much of the progress summarized here reflects sustained efforts of the LungMAP Fibroblast Task Force, whose iterative, data-driven discussions were instrumental in disentangling fibroblast identities, activation states, and contextual dependencies across datasets.

A central theme of the workshop discussions was that fibroblast identity is most consistently anchored by anatomical location. Across datasets and experimental systems, fibroblasts occupying alveolar, ductal, or adventitial compartments exhibit reproducible spatial associations that provide a stabilizing reference point, even when transcriptional profiles vary. In this context, anatomically structured nomenclatures were viewed as particularly effective, as they align with the organization of the lung along proximal–distal airway, vascular, and stromal axes. Within the stromal compartment, fibroblasts can be further contextualized by their spatial relationship to neighboring structures, including support of airways, blood vessels, or the interstitial parenchyma. This spatial anchoring was emphasized as an organizing principle that remains robust across technologies and conditions and provides a common framework for comparing fibroblast populations without requiring rigid marker-based definitions. This organizational view is intended as a practical orientation rather than a definitive classification scheme.

A second major point of consensus was that fibroblast activation states should not be equated with fixed cell types. Fibroblasts respond dynamically to local cues, and injury-induced transcriptional programs often reflect functional demands rather than lineage identity. For example, fibroblasts activated in response to epithelial injury may differ from those responding to vascular injury, even when they display overlapping functional signatures. Similarly, fibroblasts associated with airway versus vascular compartments may exhibit comparable activation programs while remaining intrinsically distinct due to their anatomical location and baseline identity. Accordingly, the working group emphasized that shared transcriptional or marker signatures, often associated with functional programs, may define states that can be adopted by multiple

fibroblast populations, rather than unique cell types, and that these states may be transient or persistent depending on context.

A further point emphasized during the workshop was that biological context is often more informative than individual molecular markers when defining fibroblast populations. Many markers commonly used to annotate fibroblasts recur across anatomical locations and activation states and may reflect shared functional programs rather than unique identities. As a result, similar transcriptional signatures can arise in fibroblasts from distinct compartments or injury contexts, even when the underlying cells differ in origin, spatial relationships, or baseline function. As an example, although a major functional structure of the lung, mesothelial cells of the pleura do not show morphological features or gene expression signatures that distinguish them from serous membranes at other anatomical sites, yet they function specifically in the pleural neighborhood. With this consideration, the group favored annotation approaches that integrate molecular features with anatomical location, physiological condition, and experimental context, allowing marker expression to be interpreted considering function and environment rather than as a standalone determinant of cell identity.

Within this framework, the working group prioritized the establishment of shared reference points by focusing initially on normal adult human and mouse fibroblasts, with the intent to extend definitions to developmental, injury-associated, and disease-emergent states in a structured manner. Progress was reported for populations with relatively stable anatomical and molecular features, including pericytes, smooth muscle cells, alveolar and adventitial fibroblasts, myofibroblasts, and activated or fibrotic fibroblast states. Ongoing discussions address distinctions among myofibroblast populations, as well as several fibroblast-associated phenotypes—including subpleural fibroblasts, lipofibroblasts, and intermediate or inflammatory fibroblasts—that are currently treated as provisional categories pending further spatial, lineage, and functional validation. A more detailed synthesis of fibroblast heterogeneity across development, homeostasis, and disease is to be addressed in a companion fibroblast-focused review.

‘OMICS: The ‘Omics Working Group identified guiding principles for the LCN project with the following recommendations. At this time, it is appropriate to start with transcriptomic clustering that will then be validated and enriched by integration with other assay modalities. Multi-batch and multi-dataset integration strengthens

identification of rare cells and pathologies but recording and reporting metadata, such as outlined in Supplemental Table 2, and strict quality control should be emphasized to avoid errors in interpretation. There may be some loss of sensitivity if quality control is so strict as to discard human or disease related cell diversity, however there is a goal in reducing harm that occurs when single batch technical artifacts are mistaken for signal. The Working Group recognized the importance of identifying and reporting best practices for assays and analyses, as suggested in this document, however with an understanding that “best practice” will change with technological and analytical advances requiring frequent comparative reassessments of methods used. Especially with this moving target, cross-consortium expertise, datasets, and resources are crucial for high-quality, high-resolution annotations. As a result, it is highly likely that cellular resolution will improve in future datasets and models so regular updates of a lung cell nomenclature are expected making it critical to maintain substantial support for a repository of cell definitions, markers, phenotypes and functions (e.g., LCN REDCap), integration of existing references and newly generated datasets (e.g., ILCA) and ongoing community annotation of healthy cells and disease-associated cell clusters (e.g., CAP).

ONTOLOGY: The Cell Ontology (CL) is currently focused on the representation of healthy cell types and states. To continue to serve as a respected reference, the CL developers expect adequate experimental evidence (at least two independent forms) to support the addition of a novel cell or the addition of identifying characteristics to an existing cell type description. CL looks to the community to verify such peer-reviewed evidence through new term requests [NTR] entered as issues in the [GitHub repository](#). Cell Ontology curators review these NTR and determine whether the existence of the cell type, or its new characteristic, is indeed supported by experimental evidence in the publication before producing textual definitions, equivalence logical axioms, descriptions, synonyms, database cross-references, taxon restrictions, and marker set confidence values using different (complementary) approaches to validate cell type identification, incorporating this new information in the subsequent ontology release. (42-44) The Brain Initiative Cell Atlas Network (BICAN), in collaboration with CL developers, demonstrated procedures in development of an extensive common cell-type nomenclature (CCN). (45) Other familiar examples of standardization with assignment of unique, meaningful names and symbols for genes and proteins, upon which cell marker identification rests, include the Human Gene Organization [HUGO Gene Nomenclature](#) (HGNC) and [UniProt](#)

that enhance non-ambiguous scientific communication, cataloging synonyms and linking to extensive genomic, clinical, and proteomic data on accessible databases.

These Working Group Task Force-led workshop discussions strengthened the commitment to the consensus lung cell nomenclature project and identified knowledge gaps that impede reaching consensus in cellular identities, including inconsistencies between modern scRNAseq markers, those reported historically, and what different groups consider to be cell type markers in this context. The activity led to a consensus proposal for further consideration: to continue to anchor on comprehensive transcriptome data with its strengths of high dimensionality and scalability, and that “a cell type marker gene [is one that] is selectively expressed, positively or negatively, above or below a determined threshold, in cells of a given type and that, used alone or in combination with other genes as canonical characteristics, reliably classify the cell type”. At several points in the process, it was expressed that recognition of a cell type requires replication across more than one independent study.

The workflow for creation and maintenance of the LCN is diagramed in Figure 3. The stimulus for the Project was the publication of multiple single cell references (1, 2, 5, 19) with inconsistent cell and marker calling despite reliance on similar data. The cell types identified in the references with highest impact at the time were used to start the LCN database with the decisions as outlined in Figure 3. Subject matter experts were engaged and an AI augmented literature search for “foundational, field-defining” lung cell biology papers emphasizing establishment of core lung cell types, ultrastructure, surfactant biology, regeneration and quantitative anatomy dating from the late 1950’s to 2026 yielded over 300 high impact publications by past and present leaders in the field (8, 10, 13, 14, 17, 29, 30, 46-61) including impactful review articles and standard anatomy texts (31, 62-78). These references were augmented by additional cell annotations from recent healthy human lung transcriptomics data to populate the LCN database. Open discussion in Working groups, ATS and NIH Workshops, as well as searching of Uberon and Cell Ontology for term matches, near matches and missing, helped to identify literature based “knowns”, “unknowns”, errors and uncertainties that were added to the LCN Cell Cards to populate the descriptive features of the database as outlined in Figure 2 and Figure 3. We focused on cells from healthy homeostatic lungs, characterized in at least 2 publications, with specific gene and protein markers, and vetted by working group subject matter experts. The requirement

to prove a cell type is a resident cell in healthy lung has been relaxed as it is a difficult bar to maintain without excluding cells of interest. The intention to work toward representing cell types and states that occur in disease will also require this flexibility in the nomenclature.

### **Current and Evolving Consensus Lung Cell Nomenclature**

This Workshop Report delivers a unified taxonomy and nomenclature for healthy adult human lower respiratory tract lung cells based on consensus in the community with a foundation in classical anatomy and histology. It provides a methodology for future integrated iterations at finer resolution to identify specific lung cell subtypes and states by both -omics content and anatomical location, using machine-readable, structured terminology with resolvable and persistent IDs that allow associated metadata (labels, synonyms, definitions, descriptions, and references) and relationships to evolve, while minimizing ambiguity and information loss. The LCN incorporates legacy knowledge with contemporary insights and novel data to link lung cell research to clinical pathology, disease, and therapeutic advances through common nomenclature.

This consensus document and the on-line resources presented are focused on cell types in healthy, adult human lung and use of the ontologies, Uberon and Cell Ontology, to document anatomical structure with cellular relationships. We have defined cell types, provisional cells, and cell states. We acknowledge and encourage that cell identities uniquely observed in lung development, maturation and/or disease should be added in subsequent iterations. Non-canonical and transitional cells are being identified even in healthy organs. For tracking and identification of knowledge gaps, we encourage addition of such cells to the LCN as provisional until sufficient defining data is available. To promote further investigation, the LCN will accommodate recording of cellular changes that occur in disease and exposure while reinforcing that they be distinguished from assay artifacts.

The [LCN webpage](#) provides an option for nominations for database additions and revisions of cell identities and a standard for review. Best practices for adding new terms to nomenclature and ontologies include: 1) a unique scientific name chosen to avoid confusion caused by colloquial names, 2) a list of synonyms aligning the cell identity to the literature, 3) use of words accepted and understood by the global scientific community, 4) ideally be clear, relatively short, easy to pronounce and descriptive, providing a rapid

indication of distinguishing characteristics, location, or function of the entity and 5) not be changed without compelling scientific reason typically after review by experts governed by specific, documented codes and guidelines to ensure consistency over time. Nomenclature is often linked to a classification system and names should have consistency with standard hierarchy or classification systems.

Entries for LCN require a term name for the cell type, a clear and concise textual definition from a domain expert that reflects the distinguishing characteristics, location, or function of the cell type, and a Parent term(s) to identify where the new term fits into the existing hierarchy (parent or subclass relation). The LCN dendrogram has shortened terms since the branching tree structure provides location, however each tree term links to a full NCBO ontological entry and, where needed, synonyms maintain legacy knowledge and latitude in naming conventions (Figure 2). When a cell type is added to the database should be accompanied by data and descriptions that clarify its origin, occurrence in the healthy lung and any known relationship to disease. Links to primary literature and reliable resources that provide evidence for the existence and characteristics of the cell type should be provided (e.g. PMID or DOI for data and image archives). The ontology system presented here chooses to refer to less well supported cell types and states as “provisional” to allow their recognition while awaiting additional investigation into source, phenotype, and function.

At this time, we introduce the LCN database content and harmonized nomenclature on canonical lung cell types. Figure 1 is the current dendrogram made accessible as a live document by the LCN. It recognizes that distinguishing cells by location, neighborhoods and niches is a priority and will become increasingly reliable and useful as highly multiplexed spatial data clarify if molecular distinctions exist. The Working Groups have reviewed each cell by superclass and the dendrogram, but data presentations and discussions continue for finer resolution and phenotyping of provisional and canonical cell terms. Figure 2 represents an example of the webpage with browsable information for the selected “alveolar type 2 cell”. Note that multiple marker sets linked to reference papers and digital resources are curated and links to photomicroscopy images are provided, Figure 2B and 2C respectively. Figures 2A, E and F, demonstrate LCN live weblinks to the database that reflect updates occurring as the work of CAP and the LCN Working Groups proceed and integrate.

Emerging areas will further enhance cell-type resolution of the lung atlases including epigenomic assays (e.g. chromatin accessibility + regulatory protein + RNA synchrony), integration of multi-molecular

assessments (protein, lipids, metabolites), high-throughput spatial assays such as spatial transcriptomics, antibody based highly multiplexed immunofluorescence, generation of true 3D and reconstructed serial section “2.5D” maps, and mass spectroscopy mediated imaging assays, as well as AI-assisted data annotation, *in silico* hypothesis generation and testing. These and accompanying computational methods-in-development will further refine cell identity, state, and function mapping across tissues and diseases for use in iterative revisions of any lung cell nomenclature and atlas effort. A useful future direction, for example, may be probing if and how homeostatic spatial location, as well as “proximalization” and “distalization” noted in disease, motivates the existence of cell identities confined to a specific location in the organ.

A formal plan for maintaining the LCN and disseminating major updates on consensus and ongoing debates for cell type/state to the broad lung research community is under development and summarized in the iterative cycle in Figure 3. The community is encouraged to provide feedback at any time through links on the LCN webpage, readily accessible during the use of the resource. The current and evolving consensus nomenclature and cell annotations will be made available through our described interactive, CAP and LCN REDCap platforms. CAP provides an interactive platform for data-driven discussions amongst experts and working groups, while the LCN database and LungMAP CellCards and LGEA repositories store the results, including images, in a compendium or condensed encyclopedia, while both support maintenance of alignment to the anatomical, cell and gene ontologies and platforms for addition of protein and other cell defining features. More granular cell-type consensus labels are also expected to be derived from the ILCA workflow and other ongoing efforts and will feed into the LCN database. Natural language interfaces like Lung Chat, a new multi-agent large-language model-based platform enables real-time interaction with single-cell and spatial transcriptomics data and nomenclature and will likely hasten discovery. In synchrony with term additions to the LCN database, additions to the Cell Ontology (CL) and Uberon Anatomy Ontology will be requested as needed. Subject matter experts are highly encouraged to contribute and assist with iterative revision of this work by accessing the Cell Annotation Platform (CAP), joining a Cell Type Working Group, or sending comments to the contact authors of this Report. The subject matter Working Groups have been asked to meet on an ad hoc basis to update nomenclature, cell markers and characteristics identified by different modalities and contexts. In addition, the working group leadership will meet approximately quarterly to review

the feedback from the interactive platforms, to curate the nomenclature and data base, and to prepare for yearly workshops. We are exploring venues and support for yearly online workshops, led by the domain experts (the working groups) and open to the community for active participation. These meetings will be virtual or hybrid and provide forums for review of recent progress in the field, including novel datasets or integrated atlases of special interest. They will provide a structured review of the most recent version of the consensus nomenclature as well as on the feedback collected on the interactive platform. The outcomes of these meetings and working group efforts will be reflected in the on-line LCN website and reported annually or bi-annually through update papers in relevant journals. We anticipate establishing a core committee to champion and approve expansion of the Lung and Airway nomenclature and ontologies and to maintain FAIR principles for the taxonomy established.

## Summary

The C3LCN ATS Assembly Project provides guidelines for single cell / single nucleus transcriptomic data generation, analysis, and publication as well as a consensus recommended nomenclature to foster a common terminology in pulmonary biology. Several data and database resources are identified. For the future work in the field, we highly encourage: a) use of standard ontologies, including correction and expansion as needed, Uberon for tissue and organ structures, and Cell Ontology (CL) for cell identities, b) refinement of annotation to be based on integrated reference datasets and community input, c) investigative community work with ontology curators to assign definitions, relationships, and attributions, d) journal and data archive requirements to use ontologies with persistent IDs, cell names, synonyms and versioning best practices, in order to reduce ambiguity, more rapidly identify and fill knowledge gaps, and as an official route to bring the identity of “new” cells to the research community, and e) establishment of support and schedule of C3LCN reviews and revisions, with provision of community updates.

## Funding and Support Acknowledgements

We gratefully acknowledge the support provided by the American Thoracic Society Assembly Project infrastructure, members of the ATS Respiratory Cell and Molecular Biology, and Respiratory Structure and Function Assemblies, and the Section on Genetics and Genomics, the Developmental Lung Molecular Atlas Program (LungMAP) and the Lung Bionetwork of the Human Cell Atlas Consortium (HCA-Lung) and all participants in the LCN Task Force Working Groups. Funding was provided by NIH U01HL148861 LungMAP Biorepository for Investigation of Diseases of the Lung (BRINDL) to G.P., R.M., G.D.; NIH Common Fund U54HL165443 “The Human Lung Biomolecular Multi-Scale Atlas Program Tissue Mapping Center (HuBMAP-Lung)” G.P., R.M., T.E.D., X.S., G.D.; U54HL165443-S2 Administrative Supplement HuBMAP-Lung: Lung Cell Nomenclature Project and MOSDAP Data Management to G.P., J.F., K.B., X.S.; NIH-NCATS UL1 TR002001 University of Rochester CTSA award; NIH 1U01HL175452 LungMAP Cross-Disease Multi-Modality Mapping of the Human Lung to X.S., T.E.D., E.S.; NIH U01HL175383 A Mechanistic 3D AI-Powered Map of Genetic chILD to K.A.W-B.; European Union H2020 grant agreement no. 874656 (discovAIR) and Horizon Europe research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 101169403 (RESPIRE-EXCEL) to MCN.; The Chan Zuckerberg Initiative, LLC Seed Network grant CZF2019-002438 “Lung Cell Atlas 1.0” to MCN.; Schmidt Futures for the Cell Annotation Platform (CAP) to A.C.V.; The LCN database is collected and managed using REDCap electronic data capture tools (27, 57) hosted at Cincinnati Children’s Hospital Medical Center 1UM1TR005265. This research was also supported in part by the Intramural Research Program of the National Institutes of Health (NIH). The contributions of the NIH authors are considered Works of the United States Government. The findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services. This document was funded by the American Thoracic Society.

### Subcommittee

This official workshop report was prepared by an *ad hoc* subcommittee of the ATS Assembly on Respiratory Cell and Molecular Biology.

Members of the subcommittee are as follows:

Gloria S. Pryhuber, M.D. (*Co-Chair*)

Martijn Nawijn, Ph.D. (*Co-Chair*)

Denise Al Alam, Ph.D.

Bruce Aronow, Ph.D.

Martin Banchero, M.S.

Nicholas Banovich, Ph.D.

Pascal Barbry, Ph.D.

Maria Basil, M.D., Ph.D.

Kevin Burns, Ph.D.

Janette Burgess, Ph.D.

Wellington V. Cardoso, M.D.

Jeremy Clair, Ph.D.

Rachel Clifford, Ph.D.

Soula Danopoulos, Ph.D.

Gail Deutsch, M.D.

Thu Elizabeth Duong, M.D.

Joshua Fortriede, Ph.D.

David Frank, M.D., Ph.D.

Reinoud Gosens, Ph.D.

Mingxia Gu, Ph.D.

Minzhe Guo, Ph.D.

James Hagood, M.D.

Raphael Kfuri-Rubens, M.D., M.Phil.

Gerard H. Koppelman, M.D., Ph.D.

Stephanie Krick, M.D., Ph.D.

Jonathan A. Kropski, M.D.

Qing Sara Lin, Ph.D.

Clare Lloyd, Ph.D.

Malte D. Luecken, Ph.D.

Susan Majka, Ph.D.

Kerstin Meyer, Ph.D.

Alexander Misharin, M.D., Ph.D.

Ravi S. Misra, Ph.D.

Ana Mora, M.D.

Enid Neptune, M.D.

Terren K. Niethamer, Ph.D.

Kenichi Okuda, M.D., Ph.D.

David Osumi-Sutherland, Ph.D.

Anna Karina Perl, Ph.D.

Joseph D. Planer, M.D.

Aleix Puig-Barbé, Ph.D.

Ellen M. Quardokus, M.S.

Rod Rahimi, M.D., Ph.D.

Jayaraj Rajagopal, M.D.

Scott Randell, Ph.D.

Elizabeth Redente, Ph.D.  
 Eniko Sajti, M.D., Ph.D.  
 Nathan Salomonis, Ph.D.  
 Christos Samakovlis, Ph.D.  
 Richard Scheuermann, Ph.D.  
 Janine Schniering, Ph.D.  
 Xin Sun, Ph.D.  
 Purushothama Tata, Ph.D.  
 Alexandra-Chloé Villani, Ph.D.  
 Matthew S. Walters, Ph.D.  
 Kathryn A. Wikenheiser-Brokamp, M.D., Ph.D.  
 Yan Xu, Ph.D.  
 Laure-Emmanuelle Zaragosi, Ph.D.

### Conflicts of interest

G.S.P. spouse/partner holds stock in AbbVie Biotherapeutics, Clearside Biomedical, CRISPR Therapeutics, GE Healthcare, Moderna, Novavax, Sorrento; received research support from National Institutes of Health. B.A. served as a consultant for Nexstone Immunology. G.D. served as a consultant for 4DMT. G.H.K. served as a consultant for AstraZeneca, PURE IMS; served as a speaker for AstraZeneca, Boehringer Ingelheim, Sanofi; received research support from European Union, Growth Fund (OMBION), Netherlands Lung Foundation, Teva, Ubbo Emmius Foundation, Vertex, ZON MW (Netherlands Organization for Scientific Research). J.A.K. served as a consultant for Arda Therapeutics; holds stock option in APIE Therapeutics; received research support from Boehringer Ingelheim, Bristol-Myers Squibb, Department of Defense, Department of Veterans Affairs, National Heart, Lung, and Blood Institute, National Institutes of Health. M.D.L. served as a consultant for Catalym; received research support from Chan Zuckerberg Initiative Foundation. K.M. employed by GlaxoSmithKline; received research support from Wellcome Trust. E.S. received research support from National Heart, Lung, and Blood Institute. P.R.T holds patents for systems and methods for lung cell expansion and differentiation, and lung fibrosis model and methods of using the same. A.C.V. served on an advisory committee for Bristol-Myers Squibb; spouse/partner holds stock option for 10x Genomics. K.A.W.B. received research support from LAM Foundation, National Heart, Lung, and Blood Institute. M.C.N. served on an advisory committee for AstraZeneca, GlaxoSmithKline; received research support from AstraZeneca, Chan Zuckerberg Initiative, Dutch Research Council, European Commission, GlaxoSmithKline, Ministry of Economic Affairs, National Growth Fund (Netherlands), Netherlands Lung Foundation. M.B., K.B., W.V.C., G.C., T.E.D., J.F., R.K.R., S.L., S.M., R.S.M., T.K.N., A.K.P., J.R., N.S., R.H.S., X.S., M.S.W., L.E.Z. reported no commercial or relevant non-commercial interests from ineligible companies.

### References

1. Guo M, Morley MP, Jiang C, Wu Y, Li G, Du Y, et al. Guided construction of single cell reference for human and mouse lung. *Nat Commun.* 2023;14(1):4566.
2. Sikkema L, Ramirez-Suastegui C, Strobl DC, Gillett TE, Zappia L, Madisson E, et al. An integrated cell atlas of the lung in health and disease. *Nat Med.* 2023;29(6):1563-77.
3. Taxonomy C. The Role of Ontologies in Artificial Intelligence and Machine Learning [Available from: [https://taxonomy.cloud/article/The\\_role\\_of\\_ontologies\\_in\\_artificial\\_intelligence\\_and\\_machine\\_learning.html](https://taxonomy.cloud/article/The_role_of_ontologies_in_artificial_intelligence_and_machine_learning.html)].
4. Pan H, Deutsch GH, Wert SE, Ontology S, Consortium NMAoLDP. Comprehensive anatomic ontologies for lung development: A comparison of alveolar formation and maturation within mouse and human lung. *J Biomed Semantics.* 2019;10(1):18.
5. Sun X, Perl AK, Li R, Bell SM, Sajti E, Kalinichenko VV, et al. A census of the lung: CellCards from LungMAP. *Dev Cell.* 2022;57(1):112-45 e2.
6. Schiller HB, Montoro DT, Simon LM, Rawlins EL, Meyer KB, Strunz M, et al. The Human Lung Cell Atlas: A High-Resolution Reference Map of the Human Lung in Health and Disease. *Am J Respir Cell Mol Biol.* 2019;61(1):31-41.

7. Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, et al. A cellular census of human lungs identifies novel cell states in health and in asthma. *Nat Med*. 2019;25(7):1153-63.
8. Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature*. 2020;587(7835):619-25.
9. Deprez M, Zaragosi LE, Truchi M, Becavin C, Ruiz Garcia S, Arguel MJ, et al. A Single-Cell Atlas of the Human Healthy Airways. *Am J Respir Crit Care Med*. 2020;202(12):1636-45.
10. Habermann AC, Gutierrez AJ, Bui LT, Yahn SL, Winters NI, Calvi CL, et al. Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv*. 2020;6(28):eaba1972.
11. Morse C, Tabib T, Sembrat J, Buschur KL, Bittar HT, Valenzi E, et al. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur Respir J*. 2019;54(2).
12. Madissoon E, Wilbrey-Clark A, Miragaia RJ, Saeb-Parsy K, Mahbubani KT, Georgakopoulos N, et al. scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol*. 2019;21(1):1.
13. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am J Respir Crit Care Med*. 2019;199(12):1517-36.
14. Goldfarbmuren KC, Jackson ND, Sajuthi SP, Dyjack N, Li KS, Rios CL, et al. Dissecting the cellular specificity of smoking effects and reconstructing lineages in the human airway epithelium. *Nat Commun*. 2020;11(1):2485.
15. Bharat A, Querrey M, Markov NS, Kim S, Kurihara C, Garza-Castillon R, et al. Lung transplantation for patients with severe COVID-19. *Sci Transl Med*. 2020;12(574).
16. Natri HM, Del Azodi CB, Peter L, Taylor CJ, Chugh S, Kendle R, et al. Cell-type-specific and disease-associated expression quantitative trait loci in the human lung. *Nat Genet*. 2024;56(4):595-604.
17. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv*. 2020;6(28):eaba1983.
18. Wang A, Chiou J, Poirion OB, Buchanan J, Valdez MJ, Verheyden JM, et al. Single-cell multiomic profiling of human lungs reveals cell-type-specific and age-dynamic control of SARS-CoV2 host genes. *Elife*. 2020;9.
19. Pryhuber G, Sun X. Human Reference Atlas HuBMAP Anatomical Structures, cell Types plus Biomarkers (ASCT+B) Tables. Lung v1.0 2021 [December 31, 2025]. 03/12/2021:[Available from: <https://apps.humanatlas.io/kg-explorer/asct-b/lung/v1.0>].
20. Borner K, Teichmann SA, Quandt EM, Gee JC, Browne K, Osumi-Sutherland D, et al. Anatomical structures, cell types and biomarkers of the Human Reference Atlas. *Nat Cell Biol*. 2021;23(11):1117-28.
21. Borner K, Blood PD, Silverstein JC, Ruffalo M, Satija R, Teichmann SA, et al. Human BioMolecular Atlas Program (HuBMAP): 3D Human Reference Atlas construction and usage. *Nat Methods*. 2025;22(4):845-60.
22. Federative International Committee on Anatomical Terminology. Terminologia histologica : international terms for human cytology and histology. Baltimore: Lippincott Williams & Wilkins; 2008. xii, 207 p. p.
23. Terminology FIPfA. FIPAT. Terminologia Anatomica. 2019. Available from: FIPAT.library.dal.ca.
24. Mungall CJ, Torniai C, Gkoutos GV, Lewis SE, Haendel MA. Uberon, an integrative multi-species anatomy ontology. *Genome Biol*. 2012;13(1):R5.
25. Diehl AD, Meehan TF, Bradford YM, Brush MH, Dahdul WM, Dougall DS, et al. The Cell Ontology 2016: enhanced content, modularization, and ontology interoperability. *J Biomed Semantics*. 2016;7(1):44.
26. Tan SZK, Puig-Barbe A, Goutte-Gattat D, Eastwood C, Aevermann B, Avola A, et al. The Cell Ontology in the age of single-cell omics. *ArXiv*. 2025;[Preprint]. 2025 Jun 17:arXiv:2506.10037v2.
27. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-81.
28. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.

29. Basil MC, Cardenas-Diaz FL, Kathiriya JJ, Morley MP, Carl J, Brumwell AN, et al. Human distal airways contain a multipotent secretory cell that can regenerate alveoli. *Nature*. 2022;604(7904):120-6.
30. Kadur Lakshminarasimha Murthy P, Sontake V, Tata A, Kobayashi Y, Macadlo L, Okuda K, et al. Human distal lung maps and lineage hierarchies reveal a bipotent progenitor. *Nature*. 2022;604(7904):111-9.
31. Rustam S, Hu Y, Mahjour SB, Rendeiro AF, Ravichandran H, Urso A, et al. A Unique Cellular Organization of Human Distal Airways and Its Disarray in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med*. 2023;207(9):1171-82.
32. Du Y, Clair GC, Al Alam D, Danopoulos S, Schnell D, Kitzmiller JA, et al. Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung. *Am J Physiol Lung Cell Mol Physiol*. 2019;317(3):L347-L60.
33. Mullan KA, Valkiers S, de Vrij N, Li C, Verbandt S, Pu T, et al. Where single-cell transcriptomics fails T cells: The misuse of unsupervised clustering for T-cell annotation. *Immunoinformatics*. 2025;20.
34. Beausoleil S, Ariss M, Huang L, Ding X, Sheth S, Levy T, et al. InTraSeq: A Multimodal Assay that Uncovers New Single-Cell Biology and Regulatory Mechanisms. *Res Sq*. 2024;[Preprint]. 2024 Dec 9:rs.3.rs-5284652. doi: 10.21203/rs.3.rs-5284652/v1.
35. Li J, Zhang Y, Yang C, Rong R. Discrepant mRNA and Protein Expression in Immune Cells. *Curr Genomics*. 2020;21(8):560-3.
36. Matsumoto R, Gray J, Rybkina K, Oppenheimer H, Levy L, Friedman LM, et al. Induction of bronchus-associated lymphoid tissue is an early life adaptation for promoting human B cell immunity. *Nat Immunol*. 2023;24(8):1370-81.
37. Masopust D, Awasthi A, Bosselut R, Brooks DG, Buggert M, Chamoto K, et al. Guidelines for T cell nomenclature. *Nat Rev Immunol*. 2026;26(4):298-313.
38. Su Y, Chen D, Yuan D, Lausted C, Choi J, Dai CL, et al. Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. *Cell*. 2020;183(6):1479-95 e20.
39. Ren X, Wen W, Fan X, Hou W, Su B, Cai P, et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell*. 2021;184(7):1895-913 e19.
40. Neumark N, Cosme C, Jr., Rose KA, Kaminski N. The Idiopathic Pulmonary Fibrosis Cell Atlas. *Am J Physiol Lung Cell Mol Physiol*. 2020;319(6):L887-L93.
41. Sauler M, McDonough JE, Adams TS, Kothapalli N, Barnthaler T, Werder RB, et al. Characterization of the COPD alveolar niche using single-cell RNA sequencing. *Nat Commun*. 2022;13(1):494.
42. Liu A, Peng B, Pankajam AV, Duong TE, Pryhuber G, Scheuermann RH, et al. Discovery of optimal cell type classification marker genes from single cell RNA sequencing data. *BMC Methods*. 2024;1.
43. Hu J, Peng B, Pankajam AV, Xu B, Deshpande VA, Bueckle A, et al. Benchmarking single cell transcriptome matching methods for incremental growth of reference atlases. *bioRxiv*. 2025;[Preprint]. 2025 Apr 16:2025.04.10.648034.
44. Xu C, Prete M, Webb S, Jardine L, Stewart BJ, Hoo R, et al. Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell*. 2023;186(26):5876-91 e20.
45. Miller JA, Gouwens NW, Tasic B, Collman F, van Velthoven CT, Bakken TE, et al. Common cell type nomenclature for the mammalian brain. *Elife*. 2020;9.
46. Weibel ER. Lung morphometry: the link between structure and function. *Cell Tissue Res*. 2017;367(3):413-26.
47. Richard JM, Williams MC. Type II alveolar cell. Defender of the alveolus. *Am Rev Respir Dis*. 1977.
48. Avery ME, Mead J. Surface properties in relation to atelectasis and hyaline membrane disease. *AMA J Dis Child*. 1959;97(5, Part 1):517-23.
49. Widdicombe JG, Pack RJ. The Clara cell. *Eur J Respir Dis*. 1982;63(3):202-20.
50. Balis JU, Conen PE. The Role of Alveolar Inclusion Bodies in the Developing Lung. *Lab Invest*. 1964;13:1215-29.
51. Hogg JC, Hackett TL. Structure and Function Relationships in Diseases of the Small Airways. *Ann Am Thorac Soc*. 2018;15(Suppl 1):S18-S25.
52. Crystal RG, Randell SH, Engelhardt JF, Voynow J, Sunday ME. Airway epithelial cells: current concepts and challenges. *Proc Am Thorac Soc*. 2008;5(7):772-7.
53. Stevens T, Phan S, Frid MG, Alvarez D, Herzog E, Stenmark KR. Lung vascular cell heterogeneity: endothelium, smooth muscle, and fibroblasts. *Proc Am Thorac Soc*. 2008;5(7):783-91.

54. Whitsett JA, Nogee LM, Weaver TE, Horowitz AD. Human surfactant protein B: structure, function, regulation, and genetic disease. *Physiol Rev*. 1995;75(4):749-57.
55. Giangreco A, Arwert EN, Rosewell IR, Snyder J, Watt FM, Stripp BR. Stem cells are dispensable for lung homeostasis but restore airways after injury. *Proc Natl Acad Sci U S A*. 2009;106(23):9286-91.
56. Basil MC, Katzen J, Engler AE, Guo M, Herriges MJ, Kathiriya JJ, et al. The Cellular and Physiological Basis for Lung Repair and Regeneration: Past, Present, and Future. *Cell Stem Cell*. 2020;26(4):482-502.
57. Perl AKT, Kistner LT, Shan SE, et al. Conditional recombination reveals distinct subsets of epithelial cells in trachea, bronchi, and alveoli. *Am J Respir Cell Mol Biol*. 2002.
58. Rock JR, Randell SH, Hogan BL. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. *Dis Model Mech*. 2010;3(9-10):545-56.
59. Carraro G, Mulay A, Yao C. Single-cell reconstruction of human basal cell diversity in normal and idiopathic pulmonary fibrosis lungs. *Am J Respir Crit Care Med*. 2020;202(11):1540-50.
60. Rothenberg ME. A hidden residential cell in the lung. *J Clin Invest*. 2016;126(9):3185-7.
61. Mesnil C, Raulier S, Paulissen G, Xiao X, Birrell MA, Pirotin D, et al. Lung-resident eosinophils represent a distinct regulatory eosinophil subset. *J Clin Invest*. 2016;126(9):3279-95.
62. Mason RJ. Biology of alveolar type II cells. *Respirology*. 2006;11 Suppl:S12-5.
63. Ruaro B, Salton F, Braga L, Wade B, Confalonieri P, Volpe MC, et al. The History and Mystery of Alveolar Epithelial Type II Cells: Focus on Their Physiologic and Pathologic Role in Lung. *Int J Mol Sci*. 2021;22(5).
64. Wang Y, Wang L, Ma S, Cheng L, Yu G. Repair and regeneration of the alveolar epithelium in lung injury. *FASEB J*. 2024;38(8):e23612.
65. Reynolds SD, Malkinson AM. Clara cell: progenitor for the bronchiolar epithelium. *Int J Biochem Cell Biol*. 2010;42(1):1-4.
66. Evans CE, Cober ND, Dai Z, Stewart DJ, Zhao YY. Endothelial cells in the pathogenesis of pulmonary arterial hypertension. *Eur Respir J*. 2021;58(3).
67. Hough RF, Alvira CM, Bastarache JA, Erzurum SC, Kuebler WM, Schmidt EP, et al. Studying the Pulmonary Endothelium in Health and Disease: An Official American Thoracic Society Workshop Report. *Am J Respir Cell Mol Biol*. 2024;71(4):388-406.
68. Mora Massad K, Dai Z, Petrache I, Ventetulo CE, Lahm T. Lung endothelial cell heterogeneity in health and pulmonary vascular disease. *Am J Physiol Lung Cell Mol Physiol*. 2025;328(6):L877-L84.
69. Kopf M, Schneider C, Nobs SP. The development and function of lung-resident macrophages and dendritic cells. *Nat Immunol*. 2015;16(1):36-44.
70. Joshi N, Walter JM, Misharin AV. Alveolar Macrophages. *Cell Immunol*. 2018;330:86-90.
71. Malainou C, Abidin SM, Lachmann N, Matt U, Herold S. Alveolar macrophages in tissue homeostasis, inflammation, and infection: evolving concepts of therapeutic targeting. *J Clin Invest*. 2023;133(19).
72. Williams M, Lambrecht BN, Hammad H. Division of labor between lung dendritic cells and macrophages in the defense against pulmonary infections. *Mucosal Immunol*. 2013;6(3):464-73.
73. Habel DM, Hogaboam CM. Heterogeneity of Fibroblasts and Myofibroblasts in Pulmonary Fibrosis. *Curr Pathobiol Rep*. 2017;5(2):101-10.
74. Ligresti G, Raslan AA, Hong J, Caporarello N, Confalonieri M, Huang SK. Mesenchymal cells in the Lung: Evolving concepts and their role in fibrosis. *Gene*. 2023;859:147142.
75. Ghonim MA, Boyd DF, Flerlage T, Thomas PG. Pulmonary inflammation and fibroblast immunoregulation: from bench to bedside. *J Clin Invest*. 2023;133(17).
76. Doherty DF, Roets L, Krasnodembskaya AD. The Role of Lung Resident Mesenchymal Stromal Cells in the Pathogenesis and Repair of Chronic Lung Disease. *Stem Cells*. 2023;41(5):431-43.
77. West JB. Invited review: pulmonary capillary stress failure. *J Appl Physiol* (1985). 2000;89(6):2483-9;discussion 97.
78. Standring S. *Gray's Anatomy: The Anatomical Basis of Clinical Practice*. 42nd ed. London: Elsevier; 2021.
79. Pryhuber GS. Concept map [110624gp Jamboree LungCellNomenclature] [Concept Map]. 2024 [updated November 23, 2025. Available from: [https://miro.com/app/board/uXjVLI3Gj3c=](https://miro.com/app/board/uXjVLI3Gj3c=/)].
80. Fullgrabe A, George N, Green M, Nejad P, Aronow B, Fexova SK, et al. Guidelines for reporting single-cell RNA-seq experiments. *Nat Biotechnol*. 2020;38(12):1384-6.

## FIGURE LEGENDS

**Figure 1. Lung Cell Nomenclature Database Considers Cell Type Hierarchy and Cell Location.** Found at LungMAP.net/research/cell-cards. The LCN Hierarchical Dendrogram begins with lineage, then structure/function context. Highlighting a cell type reveals the contents of the database relevant to that cell. Image provided by LungMAP.net as of April 21, 2026 with a Creative Commons (CC) BY 4.0 license, permitting broad reuse with appropriate attribution as to source.

**Figure 2. Lung Cell Nomenclature Database Webpage Access.** Components of the REDCap LCN Database are readily accessible from the Cell Cards Webpage. **A).** LCN full term, synonyms, closest Cell Ontology (CL) ID match, expected anatomical location, and LungMAP term ID. **B).** Gene and protein markers reported by References. **C).** Link to histologic and immunofluorescence image data, including antibody IDs, for the cell type, and Icons to open part represented in F, and to leave feedback or cell identification suggestions. **D).** Links to REDCap LCN Task Force Reports for Cell Names, Descriptions, Publication Markers and Consensus Markers, **E).** Additional details including morphological features, developmental origin, observed cell heterogeneity and states and identified cellular function, and **F).** Look-up data table linking to the LCN dendrogram, Cell Ontology, LCN Consensus gene markers in human and mouse and to the photomicroscopy image database for cell of choice. Images provided by LungMAP.net as of April 21, 2026 with a Creative Commons (CC) BY 4.0 license, permitting broad reuse with appropriate attribution as to source.

**Figure 3. Workflow for creation and maintenance of the Lung Cell Nomenclature (C3LCN).**

Graphics are original and created using Freeform (Version 3.5 (419.121.1), Apple.Inc.)

TABLES

| Table 1. LCN Adopted Glossary |                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Lung                          | The term “Lung” is used for this project in an extended sense to reflect the entire lower respiratory tract beginning with the trachea. Future work will address nomenclature issues for the upper and lower respiratory tract.                                                                                                                                                                                                                                         |
| Nomenclature                  | In biology, the system to ensure consistency and clarity by assigning standardized scientific names to features such as species, genes, proteins, or cell types based on agreed-upon rules. (e.g. HUGO Gene Nomenclature Committee; <a href="https://www.genenames.org">https://www.genenames.org</a> )                                                                                                                                                                 |
| Ontology                      | A structured, formal description reflecting an agreed-upon model of the key concepts in a specific domain such as cell types, anatomical structures, or biological processes, and the relationships between them. Ontologies provide a shared and standardized model of knowledge, enabling consistent annotation, comparison, and integration of data across studies, tools, and research groups.                                                                      |
| Taxonomy                      | The systematic classification of entities into hierarchical groups based on shared molecular, functional, anatomical, or other defining characteristics. Taxonomy organizes diversity into structured categories, enabling comparison, identification, and study across related groups.                                                                                                                                                                                 |
| Annotation                    | The process of assigning biological meaning to genomic and transcriptomic data. In cell atlasing, assigning biological identities to clusters of cells using identifying genes, regulatory elements, functional roles and anatomical characteristics. (National Center for Biotechnology Information (NCBI); <a href="https://www.ncbi.nlm.nih.gov">https://www.ncbi.nlm.nih.gov</a> )                                                                                  |
| Cell Atlasing                 | The activity of deep molecular characterization of cell identities using single-cell transcriptome sequencing and, more recently, high resolution spatial transcript and protein assays, to create a ‘CellMap’ (1) to understand all of the cell types in a tissue, organ or organism.                                                                                                                                                                                  |
| Cell Identity                 | A unique set of characteristics (structure, function, location) mediated by gene expression programs, gene regulatory networks, epigenetic profiles and interactions with neighboring cells and environment to create a specialized cell.                                                                                                                                                                                                                               |
| Cell Super Class and Class    | A cell superclass is a base / parent class that provides common attributes and methods, while a subclass (derived/child class) inherits these features and can add its own specific ones or override existing ones, forming a hierarchical relationship optimized for code reuse and specialization. These definitions are consistent with object-oriented programming (OOP) (2)                                                                                        |
| Cell Type                     | A distinct and physiologically stable physiological cell, typically irreversible under normal conditions. Typically defined by core characteristics of gene expression, supplemented by structure, function and/or anatomical location. In a database context, a cell type is defined by evidence from multiple studies showing consistent single-cell gene or protein expression markers and, ideally, characteristic anatomical localization and phenotypic features. |
| Cell Subtype                  | A cell type identifying a category within a broader cell type, distinguished by additional stable features such as unique gene or protein expression patterns, functional properties, or anatomical localization. Cell subtypes share the core characteristics of their parent cell type but exhibit finer-grained differences that justify further classification. The term is useful in hierarchical context.                                                         |
| Cell State                    | A transient or dynamically responsive property of a cell type to external (non-cell autonomous) stimuli, resulting in altered patterns of gene expression. Cell states reflect context-dependent changes such as proliferation (“cycling”), senescence (“senescent.”), or activation states in immune cells (e.g. “activated”, “exhausted” or “tolerized”) or disease-associated cell states.                                                                           |
| Cluster Identity              | A cluster identity is a group of cells, defined via data-driven clustering of single-cell omics profiles, which are characterized by common RNA expression, and thought to have common protein abundance, epigenetic signatures, and gene regulatory networks. Cell identities are detectable at a snapshot in time in a single-cell dataset, but they may be stable or transient.                                                                                      |

|                         |                                                                                                                                                                                                                                                                                                                                     |
|-------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Canonical Cell</b>   | A cell type that is broadly recognized and accepted by the scientific community as a standard or authoritative reference, based on well-established characteristics and consistent supporting evidence.                                                                                                                             |
| <b>Provisional cell</b> | A tentatively defined cell type or state proposed based on less extensive evidence than for a canonical cell. Provisional cells reflect observations that suggest a distinct identity but lack sufficient data to meet full criteria for formal cell type designation, and therefore remain subject to revision as new data emerge. |
| <b>Consensus marker</b> | A molecular marker, such as RNA transcript or protein, that is agreed upon to identify, by its presence or absence solely or in combination with other markers, a cell type or state.                                                                                                                                               |

1. Brenner S. Nobel lecture. Nature's gift to science. Biosci Rep. 2003;23(5-6):225-37.
2. Bidgoli H. Encyclopedia of information systems. Amsterdam ; Boston: Academic Press; 2003. v. <1-4> p.

**Table 2. Major Human Anatomical Lung Structures and Cell-Type Super Classes Used as the Basis for Lung Cell Taxonomy**

| I. Major Human Anatomical Lung Structures       |                                    |                                                                                                                                                                                                                                                                                                                                                   |
|-------------------------------------------------|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Structure                                       |                                    | Definition                                                                                                                                                                                                                                                                                                                                        |
| Airway Conducting Zone                          |                                    | Airways that conduct and condition air without gas exchange                                                                                                                                                                                                                                                                                       |
| Trachea                                         |                                    | Membranous tube descending from the larynx and branching into the right and left main bronchi supported by semicircular cartilaginous rings.                                                                                                                                                                                                      |
| Bronchi                                         |                                    | Upper conducting airways that arise as left and right main bronchi from the terminus of the trachea, branching into secondary lobar, then tertiary segmental, followed by 4 <sup>th</sup> , 5 <sup>th</sup> , 6 <sup>th</sup> order subsegmental bronchi; bronchial submucosa contains cartilaginous plates, serous- and mucous-secreting glands. |
| Conducting Bronchioles                          |                                    | Conducting airway distal to bronchi and prior to respiratory bronchioles; contain neither cartilage nor mucous-secreting glands; with consecutive branching, the epithelium transitions from simple multiciliated columnar to cuboidal cells.                                                                                                     |
| Submucosal Glands                               |                                    | Mucous- and serous-secreting glands in cartilaginous (bronchial) airway walls.                                                                                                                                                                                                                                                                    |
| Respiratory Zone                                |                                    | Distal lung region specialized for gas exchange.                                                                                                                                                                                                                                                                                                  |
| Respiratory Bronchioles                         |                                    | Transitional airways with occasional alveoli in the wall; first longitudinal site of gas exchange. Found in primates, pigs, ferrets, cats, and dogs but not all species including mice.                                                                                                                                                           |
| Alveolar Ducts                                  |                                    | Transitional airways that lead into alveoli.                                                                                                                                                                                                                                                                                                      |
| Alveolar Sacs / Alveoli                         |                                    | Terminal alveolar units forming the primary gas exchange surface.                                                                                                                                                                                                                                                                                 |
| Vasculature                                     |                                    |                                                                                                                                                                                                                                                                                                                                                   |
| Pulmonary Vessels                               |                                    | Vascular system dedicated to pulmonary gas exchange.                                                                                                                                                                                                                                                                                              |
| Pulmonary Arteries                              |                                    | Delivers deoxygenated blood from the right heart to the lungs.                                                                                                                                                                                                                                                                                    |
| Pulmonary Veins                                 |                                    | Returns oxygenated blood from lungs to the left heart.                                                                                                                                                                                                                                                                                            |
| Pulmonary Microvasculature                      |                                    | Conducts and controls blood flow from pulmonary arteries to veins through regions for gas exchange (Arterioles, Venules, Capillaries).                                                                                                                                                                                                            |
| Bronchial Vessels                               |                                    | Systemic vessels supplying airways and supporting lung tissues.                                                                                                                                                                                                                                                                                   |
| Bronchial Arteries                              |                                    | Systemic vessels from aorta supplying oxygenated blood to conducting zone airways and supporting structures.                                                                                                                                                                                                                                      |
| Bronchial Veins                                 |                                    | Venous vessels returning blood from conducting zone airways and supporting structures.                                                                                                                                                                                                                                                            |
| Lymphatic Vessels                               |                                    | Specialized vessels that collect and return interstitial fluid, antigens, and macromolecules from lung tissues to the lymphatic circulation.                                                                                                                                                                                                      |
| II. Human Lung Super or Parent Class Cell-Types |                                    |                                                                                                                                                                                                                                                                                                                                                   |
| Cell-Type Superclass                            | Origin                             | Definition                                                                                                                                                                                                                                                                                                                                        |
| Epithelial Cell<br>CL_0000066                   | Endoderm                           | Polarized barrier cell; In the lung, form the lung epithelium lining airways and alveoli with cell-to-cell junctional cytoskeleton, mediating barrier, secretion, absorption, and gas exchange functions.                                                                                                                                         |
| Vascular Endothelial Cell<br>CL_0002139         | Mesoderm                           | Flattened vascular lining cells forming a thin, continuous barrier in blood and lymphatic vessels; In the lung, specialized to support gas exchange, regulate permeability, and coordinate leukocyte trafficking.                                                                                                                                 |
| Leukocyte<br>CL_0000738                         | Hematopoietic stem cell / Mesoderm | Mediate host defense, immune surveillance, and tissue homeostasis; In the lung, they may be circulating or resident.                                                                                                                                                                                                                              |
| Mesenchymal Cell<br>CL_0008019                  | Mesoderm                           | Non-polarized, precursor cell, associated with matrix, capable of migration                                                                                                                                                                                                                                                                       |

|                                          |                            |                                                                                                                                                                                                                                                                                                            |
|------------------------------------------|----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Stromal Cell</b><br>CL_0000499        |                            | Structural and regulatory loose connective tissue cell. May secrete growth factors, cytokines, extracellular matrix, and support parenchymal cell mechanics and function in lung.                                                                                                                          |
| <b>Neuroendocrine Cell</b><br>CL_0000165 | Endoderm                   | Neuron capable of some hormone secretion in response to neuronal signals; In the lung, monitor oxygen, mechanical stretch and environmental stimulants and release neuropeptides to regulate airway tone, vascular function, secretion, and reflexes; Lineage tracing confirms not of neural crest origin. |
| <b>Neurons</b><br>CL_0000540             | Neural crest /<br>Ectoderm | Basic cellular unit of nervous tissue; In the lung, intrinsic, postganglionic neurons relay neuronal signals between the central nervous system and cells of the lung                                                                                                                                      |

**Table 3. Lung Mapping Consortia**

Each Program contributes through their publications, data portals, and member engagement, as data sources for the Lung Cell Nomenclature Project. They are open-access resources for exploration.

**Molecular Atlas of Lung Development Program (LungMAP)**

<https://www.lungmap.net>

LungMAP is a U.S. multi-center consortium launched in 2014 by the National Heart, Lung, and Blood Institute (NHLBI), NIH. Now in its third phase, it aims to build an open-access, dynamic 3D molecular atlas of the human lung spanning late development through adulthood. LungMAP integrates high-resolution molecular and spatial data to define lung cell types in health and disease. Its phases have focused on (1) normal postnatal lung development, (2) pediatric lung diseases such as bronchopulmonary dysplasia, and (3) adult lung health and disease. In 2022, LungMAP published *A Census of the Lung: CellCards from LungMAP*.

**Human Cell Atlas Lung Biological Network**

<https://data.humancellatlas.org/hca-bio-networks/lung>

The Lung Biological Network is one of 18 Human Cell Atlas (HCA) Bionetworks within the global HCA consortium, founded in 2016 to map every human cell type. HCA as an open and welcoming coalition, now includes more than 3,900 members from over 1,700 institutes in 100+ countries. The Lung Network focuses on creating integrated 3D cellular atlases of the lung to lead major advances in understanding of lung biology and disease. In 2023, it published *An Integrated Cell Atlas of the Lung in Health and Disease*.

**Human BioMolecular Atlas Program (HuBMAP) - including the Lung Tissue Mapping Center**

<https://hubmapconsortium.org/>

HuBMAP is an NIH Common Fund consortium developing next-generation molecular technologies and computational tools to generate foundational 3D tissue maps across the human body. Its open access expanding portal includes data from 31 organs and more than 20 assay types. HuBMAP supports understanding of cellular relationships in tissues to promote education, exploration, diagnostic and therapeutic discovery. In 2021, it began to publish Anatomical Structures, Cell Types, and Biomarkers (ASCT+B) Tables to standardize, using established and expanding biomedical ontologies, new molecular data and subject matter experts, the identification of cells in the context of anatomical structures, setting a stage to improve annotation of experimental data, integration of data types, and comparison of healthy tissues to disease states. Furthering these efforts, HuBMAP lay the foundation for the Human Reference Atlas.

Senescence Network – Lung Component

<https://Sennetconsortium.org>

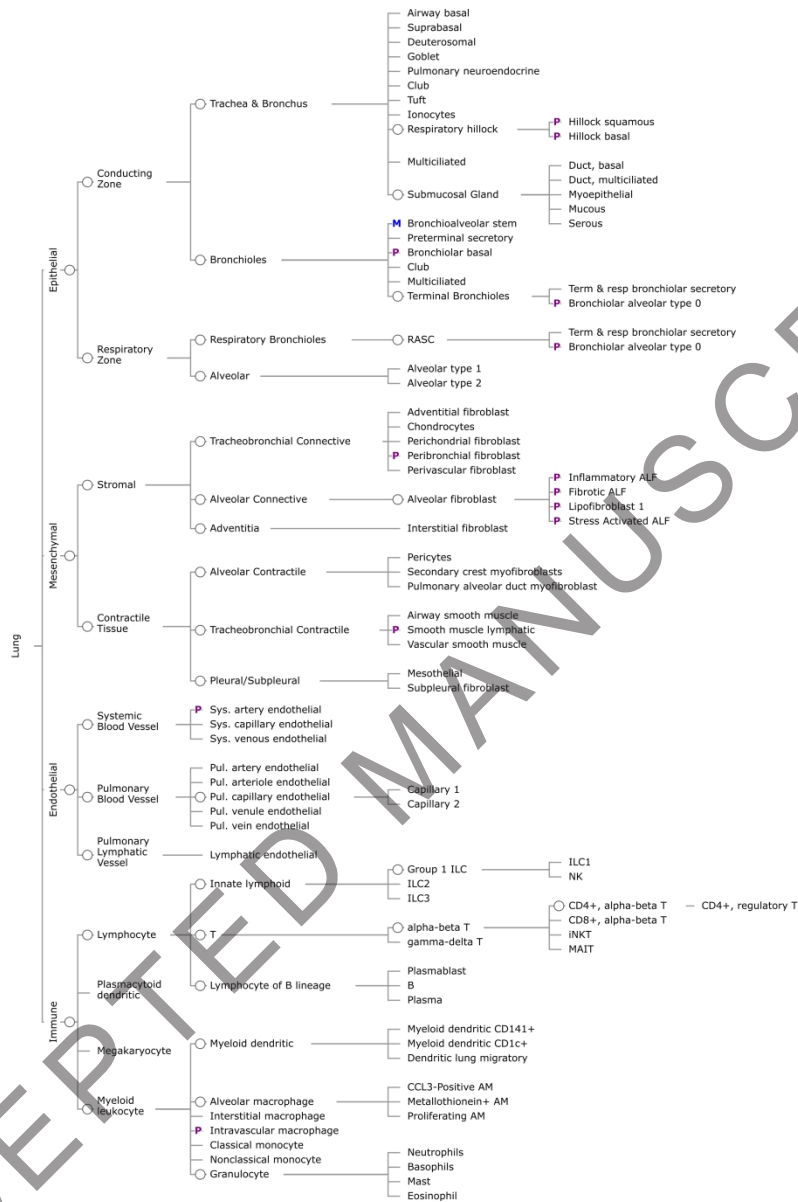
SenNet, supported by the NIH Common Fund, seeks to identify and characterize senescent cells across tissues, health states, and the lifespan. It is generating public atlases of senescent cells and their secreted factors, while developing new tools for detecting these rare cells. Building on advances from programs such as HuBMAP, SenNet also works to establish common terminology and classification systems specific for senescence biology.

Human Tumor Atlas Network (HTAN) & Lung Pre-Cancer Atlas (Lung PCA)

<https://humantumoratlas.org/>

HTAN is an NCI-funded Cancer Moonshot initiative creating 3D atlases of cancers as they progress from precancerous lesions to advanced disease. Using molecular, cellular, and morphological profiling, HTAN studies tumor initiation, progression, metastasis, treatment response, and resistance. Lung-focused efforts include the Lung Pre-Cancer Atlas, which examines early lung carcinogenesis and disease evolution.

ACCEPTED MANUSCRIPT



665x1000mm (118 x 118 DPI)

### A. Dendrogram Label

Alveolar type 2

**Synonyms** alveolar epithelial cell 2, type II alveolar cell, lung alveolar epithelial type II cell, Alveolar type 2 cells, type 2 pneumocyte, AT2, type II lung epithelial cell, AEC2

**Cell Ontology** pulmonary alveolar type 2 cell (CL:0002063)

**Cell Ontology Match Type** Exact

**Species** Human: Present  
Mouse: Present

**Location** alveolus of lung, pulmonary acinus, alveolar sac

**Term ID** LMCC0000000004

**Cell Task Force Reports**

**Cell Names**

**Cell Descriptions**

**Publication Markers**

**Consensus Markers**

### B. Details

**Descriptions**

**Reference Datasets**

**Images**

**Term:** alveolar type 2 epithelial cell

| Species | Dataset                                                                                        | Marker Type | Markers                                                           |
|---------|------------------------------------------------------------------------------------------------|-------------|-------------------------------------------------------------------|
| Human   | Cell Ref (Guo et al 2023)                                                                      | Gene        | ABCA3 LAMP3 KCNJ15 SFTPC SFTPA1 SFTPA2 SFTPB PGC NAPSA SFTPD      |
| Human   | HLCA (Sikkema et al. 2022)                                                                     | Gene        | MFSD2A C8orf44 C11orf96                                           |
| Human   | ASCT+B                                                                                         | Gene        | BMP1 SFTPA2 LAMP3 MFSD2A ABCA3                                    |
| Human   | Cell Cards (Sun et al)                                                                         | Gene        | SFTPC LAMP3                                                       |
| Human   | Cell Cards (Sun et al)                                                                         | Protein     | KCNJ15                                                            |
| Human   | Scheuermann NS Forest                                                                          | Gene        | MFSD2A C11orf96                                                   |
| Human   | Scheuermann NS Forest                                                                          | Gene        | SFTPA1                                                            |
| Human   | Duong, et al. snHLA (A Multimodal Spatial and Epigenomic Atlas of Human Adult Lung Topography) | Gene        | SFTPB SFTPC BMP1 ABCA3 LAMP3                                      |
| Human   | Duong, et al. snHLA (A Multimodal Spatial and Epigenomic Atlas of Human Adult Lung Topography) | Gene        | SFTPB SFTPC BMP1 ABCA3 LAMP3 CCDC41 LINC11-AS1                    |
| Human   | Duong, et al. snHLA (A Multimodal Spatial and Epigenomic Atlas of Human Adult Lung Topography) | Gene        | SFTPB SFTPC BMP1 ABCA3 LAMP3 LRP2 SCN1A DMBT1                     |
| Mouse   | Cell Ref (Guo et al 2023)                                                                      | Gene        | Lamp3 Abca3 Kcnj15 Sftpa1 Lyz2 Lyz1 Lys1 Sftpb Sftpc Scn7a2 Ch3f1 |
| Mouse   | Cell Cards (Sun et al)                                                                         | Gene        | Shcna_lamp3                                                       |
| Mouse   | Cell Cards (Sun et al)                                                                         | Protein     | Kcnj15                                                            |

### C. Details

**Descriptions**

**Reference Datasets**

**Images**

**Term:** alveolar type 2 epithelial cell

**Dendrogram Label** Alveolar type 2

**Morphological Features** cuboidal alveolar epithelial cells with lamellar bodies and specialized microvilli and lamellar bodies that can be defined by TEM and further confirmed with staining in IHC or IFC;

also contain evidence for a lysosome-like organelle compartment, which is acidic, and rich in phospholipid, and containing mature SP-B/proSP-C proteins.

In mouse: cuboidal alveolar epithelial cells with lamellar bodies and specialized microvilli that can be defined by TEM and lamellar bodies further confirmed with staining in IHC or IFC

also contain evidence for a lysosome-like organelle compartment, which is acidic, and rich in phospholipid, and containing mature SP-B/proSP-C proteins.

**Developmental Origin** Human: By 15 pcw the AT2 markers SFTPC and SFTPA were detected by combination of single-cell RNA and ATAC sequencing, high-throughput spatial transcriptomics, and single-cell imaging from SOX9/ID2+ multipotent progenitors [He et al. Cell 185 (2022)]

In mouse: Mouse: Specialized AT2s can be detected by *Sftpc* RNA in situ and lineage tracing at E14.5-E15.5, likely originating from Sox9/Id2+ distal tip cells [Frank, et al. 2019; Zeng, et al. 2021]. In postnatal lungs, after following injury, lost AT2 cells are primarily replaced by remaining AT2 cells. (CellCards)

**Heterogeneity and Cellular States** AT2s harbor at least one progenitor sub-lineage (WNT-responsive *Axin2*+ alveolar epithelial progenitors or AEPs) that are critical for alveolar regeneration after injury [Kashani, et al., 2018; Zacharias, et al., 2018]

Both human and murine AT2 cells can enter a transitional intermediate state that can be accelerated by exogenous injury or endogenous reprogramming. It remains to be determined whether this state is physiological, pathological, or both: HUMAN Transitional Cell State: Krt5+Krt17+ + DAPIs, PATS (i) Hubermann et al *Sci Adv* 2020; (ii) Adams, et *Sci Adv* 2020

**Function**

a) mature and functional system to produce surfactant, a protein and lipid mixture that reduces surface tension to allow lung expansion during inhalation

b) barrier function coupled with AT1 cells

c) participation in lung host defense through expression of collectins (SP-A and SP-D) and cytokines that promote control of pathogens and effector cell (inflammatory) response.

d) regenerative (stem-cell) activity

### D. Lung Cell Data Table

Clicking on the cell name will select the corresponding cell in the tree.

Search:

| Cell Name                       | Anatomy                                                              | Other names                                                                                                                                                                 | Cell Ontology                   | Consensus (Human)                         | Consensus (Mouse)                                  | Images |
|---------------------------------|----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------|-------------------------------------------|----------------------------------------------------|--------|
| alveolar type 1 epithelial cell | alveolus of lung, alveolar sac, epithelium of respiratory bronchiole | AT1, AEC1, alveolar epithelial cell 1, Alveolar type 1 cells, type 1 pneumocyte, AT1                                                                                        | pulmonary alveolar type 1 cell  | 10 Reference Datasets                     | 3 Reference Datasets                               |        |
| alveolar type 1 fibroblast cell |                                                                      | matrix fibroblast 1, Alveolar fibroblasts, lipofibroblasts, Alveolar Fibroblast 1 cells                                                                                     | alveolar type 1 fibroblast cell | SCN7A, PDGFRA, TCF21, WNT2, FGFR4, SCUBE2 | Scn7a, Pdgfra, Tcf21, Wnt2, Fgfr4, Scube2, Col12a1 |        |
| alveolar type 2 epithelial cell | alveolus of lung, pulmonary acinus, alveolar sac                     | alveolar epithelial cell 2, type II alveolar cell, lung alveolar epithelial type II cell, Alveolar type 2 cells, type 2 pneumocyte, AT2, type II lung epithelial cell, AEC2 | pulmonary alveolar type 2 cell  | 10 Reference Datasets                     | 3 Reference Datasets                               |        |

487x647mm (118 x 118 DPI)

