Machine learning analysis of the bleomycin-mouse model reveals the spatial and temporal pulmonary inflammatory fingerprint

Natalie Bordag[1], Valentina Biasin[1,2], Diana Schnoegl[1], Valzano Francesco[1], Katharina Jandl[1,3], Bence M. Nagy[1], Neha Sharma[1,4], Malgorzata Wygrecka[5], Grazyna Kwapiszewska[1,3], Leigh M. Marsh[1*]

Affiliations

[1]Ludwig Boltzmann Institute for Lung Vascular Research, Graz, Austria.

[2]Division of Endocrinology and Diabetology, Department of Internal Medicine, Medical University of Graz, Graz.

[3]Otto Loewi Research Center, Medical University of Graz, Graz, Austria.

[4]Division of Pulmonology, Department of Internal Medicine, Medical University of Graz, Graz, Austria.

[5]Department of Biochemistry, Universities of Giessen and Marburg Lung Center, Giessen, Germany. Member of German Center for Lung Research.

*to whom correspondence should be addressed:

Dr. Leigh Marsh
Leader of the Translational Platform
Ludwig Boltzmann Institute Lung Vascular Research
Center for Medical Research
Neue Stiftingtalstraße 6/VI
8010 Graz, AUSTRIA
Leigh.Marsh@lvr.lbg.ac.at

One Sentence Summary

Unbiased immunophenotyping and data modelling exposed the dynamic shifts in immune cell composition during bleomycin induced pulmonary fibrosis.

Abstract

The bleomycin mouse model is the most extensively used animal model to study pulmonary fibrosis. Despite this, the inflammatory cell kinetics and cell compartmentalisation is still incompletely understood. Here we simultaneously analysed 16 inflammatory cell populations in 303 samples and applied advanced data modelling to conclusively detail these kinetics.

Three days post-bleomycin, the inflammatory profile was typified by acute innate inflammation, pronounced neutrophilia and loss of alveolar macrophages. After 14 days, rapid responders were increasingly replaced by the adaptive immune system and monocyte-derived alveolar macrophages, which progressed till 21 days. Multi-colour imaging revealed the spatial-temporal cell distribution and the close association of T-cells with fibrotic lung tissue at later time-points.

Unbiased immunophenotyping and data modelling exposed the dynamic shifts in immune cell composition distinct for each phase of fibrosis process and defined the transition from innate to adaptive immunity marking initial lung parenchyma remodelling.

Keywords

Bleomycin, mouse, inflammatory cells, machine learning, pulmonary fibrosis

Abbreviations

BALF, bronchoalveolar lavage fluid; BH, Benjamini-Hochberg; FCM, flow cytometry; IPF, idiopathic pulmonary fibrosis; LOG_{x+1}, logarithm to the basis 10 of (x+1); ML, maximum likelihood; MVA, multivariate analysis; NLME, non-linear mixed models; OPLS-DA , orthogonal projections to latent structures discriminant analysis; PCA, principal component analysis; PF, pulmonary fibrosis; UMAP, Uniform Manifold Approximation and Projection; UVA, univariate analysis; 4RT, fourth root.

Introduction

Animal models of human disease are an invaluable tool to decipher disease relevant pathomechanisms, to discover therapeutic targets and to drive translation into clinical practice. To date, the mouse bleomycin-induced lung injury model is the most frequently used animal model to investigate pulmonary fibrosis (B Moore et al., 2013; Della Latta et al., 2015; Tashiro et al., 2017). Similar to the human situation, in mice bleomycin exposure is characterized by epithelial damage, inflammatory cell infiltration, and expansion of fibroblasts and myofibroblasts as well as ECM deposition (Biasin et al., 2020, 2017; El Agha et al., 2017; Tashiro et al., 2017; Xie et al., 2018). Although, the bleomycin model does not completely recapitulate human idiopathic pulmonary fibrosis (IPF), it still remains the most common and important animal model to study this disease.

IPF is a severe, rapidly progressing interstitial lung disease with high mortality rates and short median survival of 1.5 - 4 years (Marshall et al., 2018; Wuyts et al., 2013). IPF is characterized by extensive lung tissue scarring, limited inflammation and extracellular matrix remodelling (Meltzer and Noble, 2008). Current treatment options slow the loss of lung function, but are unable to halt or reverse disease progression (Maher and Strek, 2019). Accordingly, there is an urgent unmet clinical need for novel therapies for IPF patients. To date the aetiology and pathogenesis of IPF is still insufficiently understood; however, the role of inflammation remains undeniable yet controversial. The older concept that IPF is an inflammatory driven process has been gradually replaced by the theory of recurrent injury and aberrant repair (Selman et al., 2001; Selman and Pardo, 2002; Wuyts et al., 2013). However, multiple inflammatory cells have been implicated in disease pathogenesis, including macrophages (Misharin et al., 2017; Reyfman et al., 2019) and T cells (Todd et al., 2013), which are connected with poorer prognosis (Balestro et al., 2016).

In the bleomycin model, the early phase post bleomycin administration is characterised by acute lung injury and inflammation, which is observed to last between 1-7 days (Peng et al., 2013). This inflammatory phase is followed by active fibrosis, between 7-14 days and late fibrosis between, 21-28 days (Della Latta et al., 2015; Izbicki et al., 2002; Peng et al., 2013; Tashiro et al., 2017). As most studies have only analysed specific cell populations or time points, a comprehensive description of the inflammatory cell kinetics is still missing. For the detection and quantification of inflammatory cells, flow cytometry (FCM) is the method of choice. FCM is able to differentiate and quantify immune cell populations in unprecedented detail, not only from the circulation but also from disease relevant tissue (Marsh et al., 2018; Misharin et al., 2017; Tighe et al., 2019a). In contrast to traditional immunofluorescent staining, which generally use 1-3 markers for cell identification, flow cytometers applies multiple markers to simultaneously quantify numerous cell populations at a single cell resolution. Thus, FCM generates large quantities of complex data, where the analysis, visualization and

88 interpretation of which requires sophisticated analysis techniques, such as computational flow
89 cytometry (Saeys et al., 2016).

90 In order to conclusively detail the inflammatory cell kinetics in the bleomycin model, we here
91 assembled historical FCM data from 15 different experiments and applied advanced data modelling,
92 including univariate, multivariate and machine learning methods. We show how the combination of
93 advanced data modelling and in-depth immune profiling can detail the dramatic changes in the
94 inflammatory landscape in this model and also serves as a reference point.

Results

Pre-processing of flow cytometric data substantially improves statistical analysis performance

Intra-tracheal administration of bleomycin in mice, results in a time-dependent development of fibrosis (Figure 1AB). To comprehensively describe the inflammatory cell kinetics following bleomycin treatment, we assembled and conjointly analysed historical FCM data from 15 independent experiments, this resulted in 159 BALF and 144 lung tissue samples (Supplementary Table S1). Using standard gating strategies, a total of 16 cell populations covering the main myeloid and lymphoid cell types (Table 1) were identified (Figure 1C). The aggregation of historical experiments inherently led to an unbalanced experimental design (Supplementary Table S1), which was handled by robust statistical methods[Box 1].

Table 1. Inflammatory cell identification and corresponding markers.

					Panel I								Panel II				
Panel	**Cell type**	**Abbrievation**	**Cell population**	**SSC**	**CD45**	**SiglecF**	**CD11c**	**Gr-1**	**CD64**	**CD24**	**MHC-II**	**CD11b**	**CD3**	**CD4**	**CD8**	**CD19**	**gdTCR**
	Cell count		$CD45^+$ live cell		+												
Myeloid	Macrophages	AM	Alveolar macrophages	hi	+	+	+		+			+/-					
		MoAM	Monocyte derived macrophages		+	lo	+		+	-	+	+/-					
		IM	Interstitial macrophages		+	-	-		+	-	+	+/-					
	DCs	DC	$CD11b^+$ Dendritic cells		+			-	-	+	+	+					
	Monocytes	$Gr1^+$ MoMp	Inflammatory monocytes		+			+	+		-	+					
		$Gr1^-$ MoMp	Constituative monocytes		+			-			-	+					
	Granulocytes	EOS	Eosinophils	hi	+	+	-					+					
		PMN	Mature neutrophils	hi	+	+/-	-	+				+					
		$SiglecF^+$ PMN	$SiglecF^+$ neutrophils	hi	+	+	-	+				+					
		Immature PMN	Immature neutrophils	hi	+		-	+			-	+					
Lymphoid	B cells	$CD19^+$ B cells	B cells	lo	+								-			+	
	T cells	$CD3^+$ T cells	T cells	lo	+								+			-	
		$CD4^+$ T cells	T helper cells	lo	+								+	+	-	-	
		$CD8^+$ T cells	Cytotoxic T cells	lo	+								+	-	+	-	
		γδ T cells	γδ T cells	lo	+								+	+	-	-	+

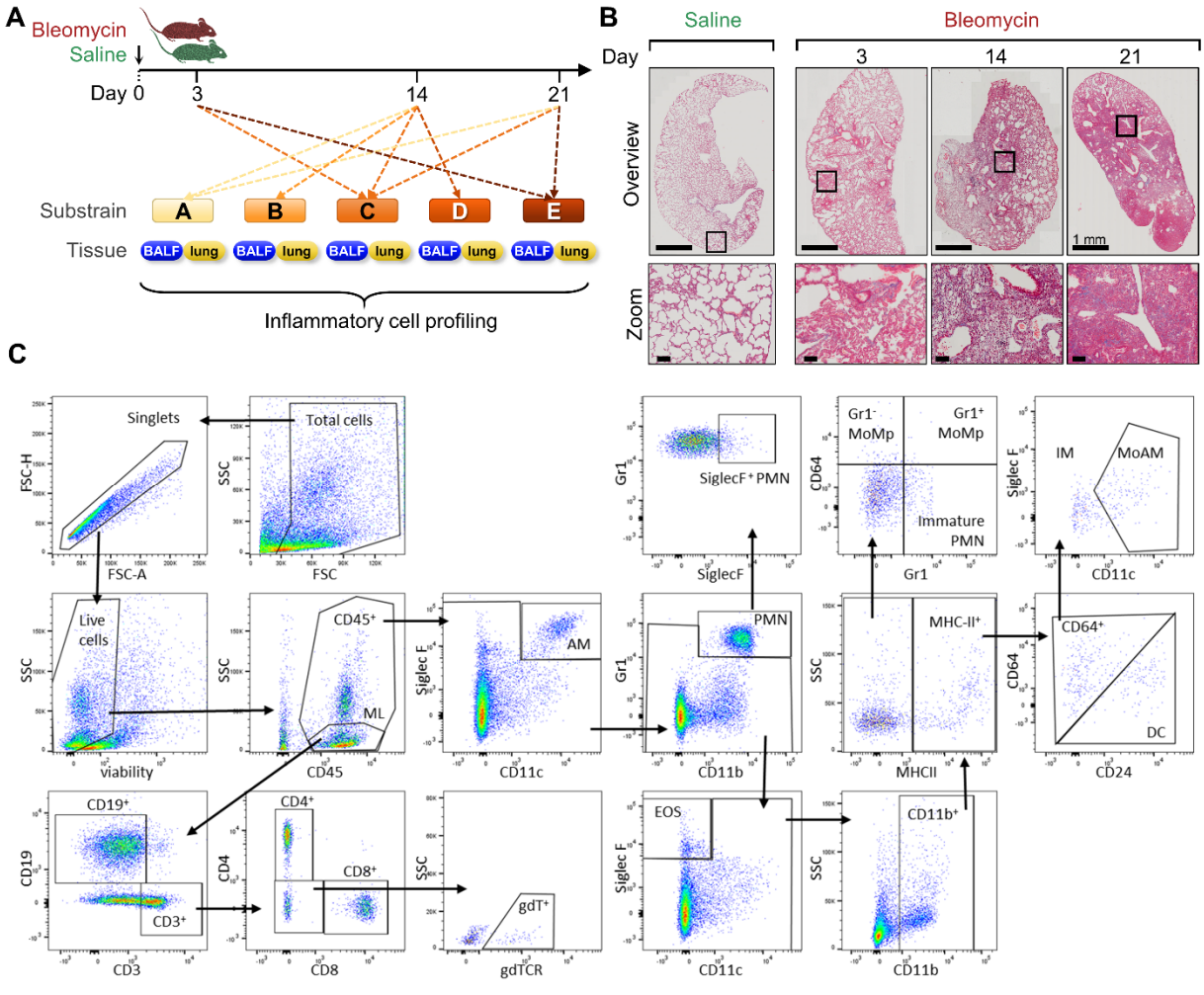

Fig. 1. Overview of study design, pathological changes and gating strategy. (A) Historical flow cytometry data from the bleomycin mouse model were pooled and collectively analysed. Samples were collected 3, 14, or 21 days post bleomycin or saline administration from the compartments BALF (159 samples) and lung tissue (144 samples). Five different C57BL/6 substrains were included. (B) Representative Masson's trichrome staining of lung sections, showing pathologic alterations in the bleomycin model. Zoomed images exemplify the increasing fibrosis accumulation from day 3 to 21 after bleomycin challenge, scale bar represents 100 µm. (C) Representative flow cytometry gating strategy. Abbreviations see Table 1.

In both tissues the distribution of all 16 analysed cell populations was significantly non-normal with a positive skew[Box 1] (Fig. 2A, Supplementary Fig. S1 and Supplementary Data 1). To improve distribution we trialled several common transformations; square root, reciprocal, Freeman Tukey, logit, LOG, LOG_{x+1} and 4RT. Only LOG, LOG_{x+1} and 4RT improved data distribution ($p_{BH}>0.05$, Supplementary Data 1). As both LOG and LOG_{x+1} gave virtually equivalent results, but as LOG_{x+1} has additionally the advantage of not introducing missing values for zero value counts, consequent analysis was performed with only LOG_{x+1} and 4RT (Fig. 2AB).

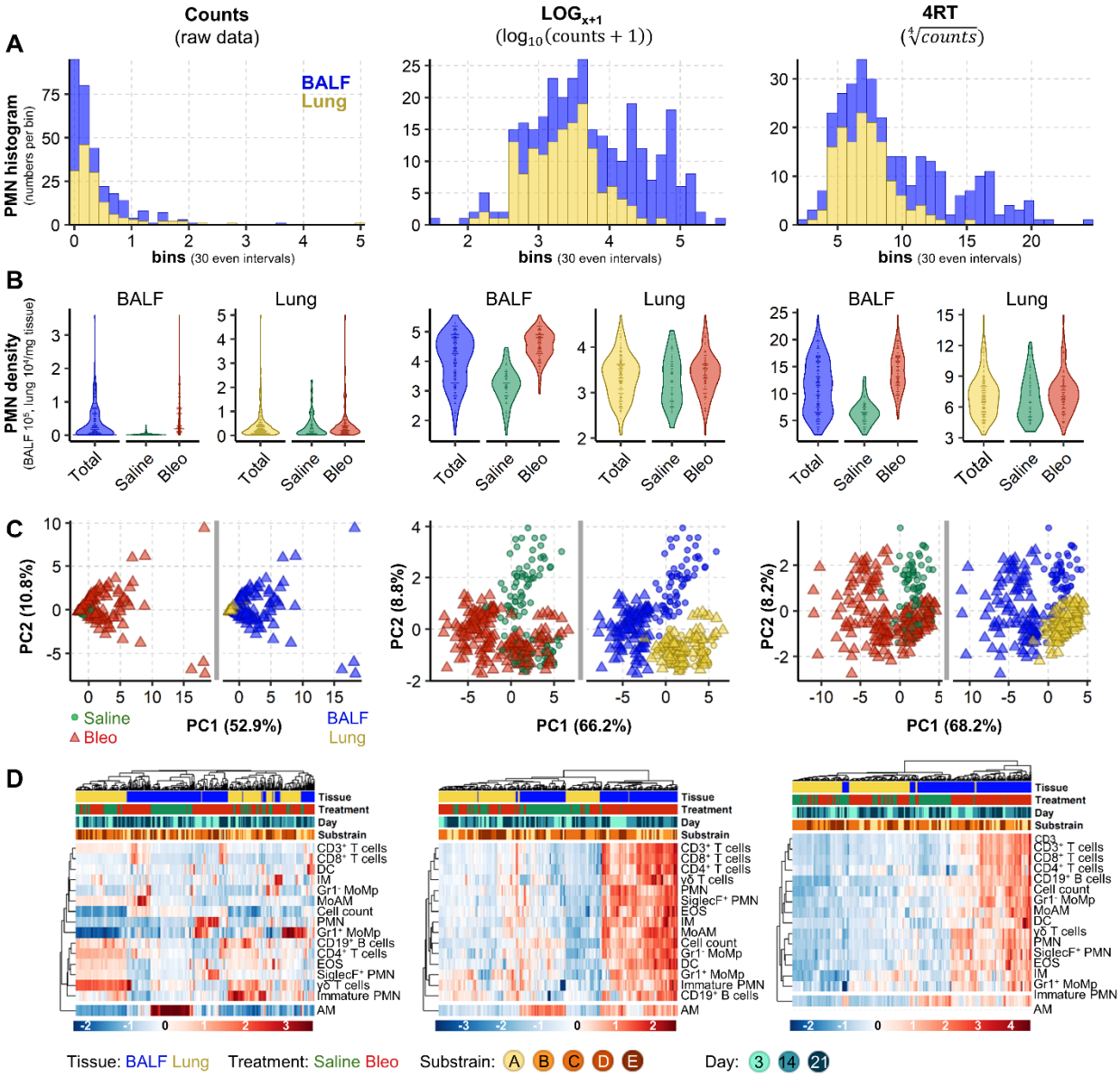

Fig. 2. Data transformation improves data distribution and analytical power. Analysis of cell count data (untransformed) or following transformation using LOG_{x+1} or 4RT (fourth root) using 159 BALF and 144 lung samples. Cell counts in BALF are 10^5 and in lung 10^4/mg tissue. Examples of data distribution of neutrophils (PMN) as one representative population in BALF and lung samples by (A) Histograms and (B) Violin plots, total represents combined saline and bleomycin samples. (C) PCA scores plots with each point representing the inflammatory cell profile (16 populations) in one sample, plots are coloured to highlight different experimental conditions. In B and C, dots represent single sample values. (D) Heatmaps with hierarchical clustering of all 16 analysed cell populations.

Bleomycin drives strong changes in the inflammatory profile

To identify global changes in the inflammatory cell profile, we first applied unsupervised principal component analysis (PCA). This method reduces dimensionality by creating new variables, which successively maximize variance and thereby aids data interpretability. Without data transformation, the scores plot was dominated by single sample differences, which obscured any experimental effects (Fig. 2C, left panel). After transformation pronounced differences in the inflammatory profile were revealed

(Fig. 2C). Both LOG_{x+1} and 4RT substantially improved the performance of the hierarchical clustering (HC), yielding clearer clustering and heatmap results (Fig. 2D). The highest influence on the inflammatory landscape was caused by the tissue compartment (BALF or lung), causing samples to separate along the first principal component (PC1). The second highest difference was caused by bleomycin, separating samples in the BALF along the second principal component (PC2; Fig. 2C, middle and right panels). Similarly, HC was first driven by the tissue compartment, followed by some weaker subclustering due to bleomycin treatment. The majority of cell populations increased after bleomycin exposure, while alveolar macrophages (AlvMp) decreased (Fig. 2D). We next utilised macroPCA, a robust PCA method able to handle and identify all possible types of data contaminations[Box 1], including strong single value or sample outliers (Hubert et al., 2019). MacroPCA results were in good agreement with PCA (Supplementary Fig. S2A), which confirmed that this dataset is free of severe outliers allowing the use of a wide variety of statistical methods (Rousseeuw and Hubert, 2018).

Box 1 | Glossary of analysis terms

robustness	is a measure for how easily outlier values distort results, e.g. • average: not robust, a single strong outlier deforms results severely • median: very robust, good results even with almost half of all values being strong outliers
unbalanced	describes unequal group sizes or missing values, methods assuming balanced groups will have missleading results
positive skew	asymmetric distribution of data with more large than small values, common in flow cytometry and many other biological measures (often because zero is the minimum, while there is no fixed maximum)
data preprocessing	preprocessing normalises data by changing all values according to one or several defined mathematical equations and can be a prerequisite for specific statistical methods Centring and scaling • cell count differences are not per se reflective of their biological importance • centring and scaling minimizes the stark differences of cell numbers between the cell populations allowing comparisons of fold changes • are vital for multivariate statistical methods, otherwise results will be dominated by cells with highest counts/highest noise Transformation • improves data distribution allowing use of more powerful statistical methods (Keene, 1995; van den Berg et al., 2006) all types of preprocessing can be combined with each other
centring	subtraction of a constant from every value (e.g. the average of each cell population)
scaling	every value is divided by a constant (e.g. the standard deviation, SD)
transformation	convert each data point by a specific, often nonlinear, but defined mathematical function (e.g. log10)
data contaminations	denotes all kinds of problematic values in the data, such as sample outliers, single value outliers or missing values
outlier	a value so different from the rest, that it could be for example an analytical error

As the strong compartment effect could mask weaker drivers that alter the inflammatory landscape, we analysed BALF and lung samples separately (Fig. 3). In the BALF, bleomycin exposure completely

altered the inflammatory landscape, separating samples along PC1 (explaining 63.9 % of the variation in the dataset). However, the bleomycin effect only accounted for 12.4 % of the variation in the lung, separating on PC2 (Supplementary Fig. S2A). Again, macroPCA gave similar results in the analysis of the separate compartments (Supplementary Fig. S2B), reconfirming absence of critical outliers. Analogous to the PCA findings, HC showed a strong clustering after bleomycin exposure in BALF, which was less clear in lung. The influence of day post-treatment and substrain (individual C57BL/6J lines) on cell population changes was less distinct, with only some indication towards a possible sub-clustering due to these factors (Fig. 3B).

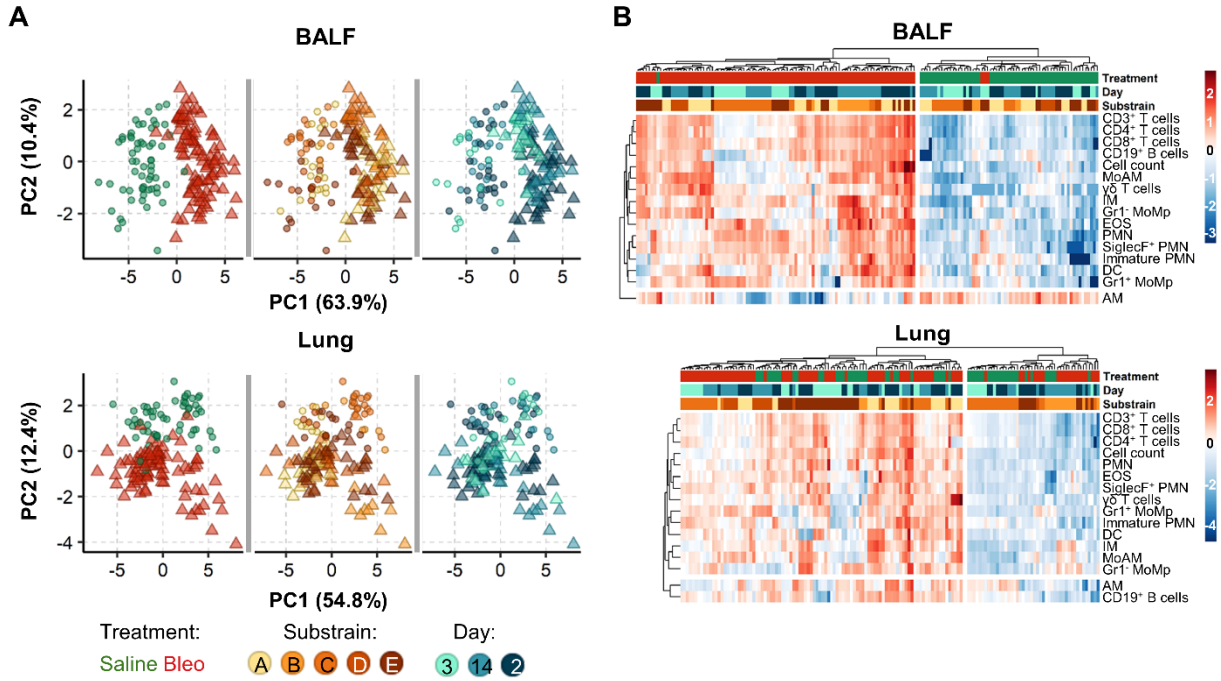

Fig. 3. Bleomycin induces stronger changes in the inflammatory profile in the BALF than the lung. The contribution of different biological factors to the inflammatory cell profile as determined by (A) PCA scores plots are coloured to highlight different experimental conditions, and (B) Heatmaps with hierarchical clustering. To aid interpretation heatmaps are split into two main clusters based on dendrogram distances. Colours and shapes represent tissue, treatment (Saline, Bleo), mouse substrain and day post treatment. Cell counts from 16 populations in 159 BALF and 144 lung samples were LOG_{x+1} transformed prior to clustering.

Modelling of inflammatory cell kinetics with univariate statistical analysis

In order to examine in depth, the potential influence of other experimental factors and to simultaneously control for the unbalanced design arising from the use of historical data, we applied non-linear mixed models (NLME, Supplementary Fig. 3). As the multivariate analysis showed a strong bleomycin effect, the fixed factor[Box 2] *Treatment* {Saline,Bleo} was included in all models[Box 2]. Other fixed factors included *Day* {3,14,21} and *Substrain* {A,B,C,D,E}. The addition of each factor, either alone or together and with or without their interaction with *Treatment*, notably improved the fit[Box 2] of all simple

179 models, increasing the goodness of fit and reducing AIC (Supplementary Fig. S3). Thus, both the *Day*
180 post bleomycin exposure and *Substrain* significantly influenced the cellular landscape.

181

Box 2 | Glossary of model terms

model	a mathematical equation to describing the relationship of measured data to biological factors • imagine you assume that the weight increases with height, than the biological factor is body height, the measured data is weight and a linear model would have the equation: *weight* = *a*·*height* + *b* • parameters are *a* and *b*, *a* – inclination (steepness of the line), *b* –intercept (*weight* at *height*=0)
fitting	finding the best parameter values in the mathematical equation of the model, parameters are optimized to bring the line/curve of the model nearest to the data, often assessed by the residuals
fixed factor	also called between-subject effect, a biological factor which (possibly) affects the outcome • height is a fixed factor in the example, gender would be another possible fixed factor
interaction	the impact of one biological factor depends on the occurrence of another biological factor • imagine to include also gender and genetics as biological factors, the effect that males are roughly 0.13 m taller than females is much smaller in achondroplasia which results in short stature
random factor	also called within-subject effect, a factor which (possibly) affects baseline level such as repeated measures from the same source or working in experimental batches • in linear model example that is to measure the height/weight yearly during adolescence while smaller
simple/mixed	simple models have no random factor, mixed models have a random factor
residuals	difference between fitted value and measured value • in linear model example that is the distance from the measured value to the line
fitted value	the value suggested by the equation for the specific value of the biological factor • in linear model example that would at a given height the weight on the line
predicted value	similar to fitted values the predicted value is suggested by the model equation, but for formerly unknown points (unknown during model fitting or not measured at all) • imagine the linear model was based on heights from 1 – 1.5 m and you want to predict the weight for 1.7 m
overfitting	the model contains more parameters than possible from the data, fails to predict new data correctly • in the example adding irrelevant factors (e.g. birthdates, house numbers, number of earrings, …) can produce perfect fits but fail to predict new values

182

183 As each independent experiment could have similarities, the experimental ID was then included as a
184 random factor (~1|Exp_ID). These mixed models significantly outperformed the aforementioned simple
185 models. Finally, complex mixed models (combining the mixed models with the interactions of
186 *Treatment* with *Substrain* or *Day*) notably outperformed all simple models (with or without
187 interactions). The most complex mixed model [*Treatment*+*Day*+*Substrain*+*Treatment:Substrain*+
188 *Treatment:Day, ~1|Exp_ID*] outperformed all other models, although more prominently in BALF than
189 in lung (Supplementary Fig. S3A).

190 As complex models risk overfitting, especially in light of the unbalanced design, we then investigated
191 model simplification. We first tested whether it was possible to create one control group of all saline
192 animals. In all mixed and complex models (i.e. with random factor *Exp_ID*) only 4 of the over 10000
193 investigated pairwise comparisons of a saline subgroup with another saline subgroup had a $p_{BH}<0.01$ in

any of the 16 cell types. This means saline treated animals were sufficiently similar to be combined into one control group. Consequently, *Treatment* and *Day* can be then merged into one fixed factor with four groups: Saline (all days) and bleomycin after days 3, 14, and 21, which was termed *SalineDay* {Saline,3,14,21}, generating the simplified model [*SalineDay*+*Substrain*] and the simplified mixed model [*SalineDay*+*Substrain~1| Exp_ID*]. The performance of the simplified mixed model was slightly lower than in the most complex mixed model, but well within the range of the other top performing mixed models (Supplementary Fig. S3B).

To compare the models in more detail we also directly compared the fitted values[Box 2] of the simplified mixed model with the most complex mixed model. The fitted values from both models strongly correlated (Pearson correlation R^2>0.96, Supplementary Fig. S3B). This underlines the validity of model simplification and that no unexpected or systematic skew was introduced. As the simplified mixed model [*SalineDay*+*Substrain~1|Exp_ID*] also gives more easily interpretable results and has a lower risk of overfitting[Box 2], it was chosen to examine the inflammatory cell kinetics underlying bleomycin mouse model.

This model was then applied to explore how individual substrains may influence the kinetics of different inflammatory cells. All mice included in this study are on the C57BL/6 background, however were obtained from different sources e.g. commercial sources (C57BL/6J, substrain A), or are the wild-type littermates from in-house breedings (substrains B-E). Although some lines were inbred for up to 15 generations, all mouse lines produced similar inflammatory responses in both lung compartments, differing only in magnitude (Supplementary Fig. S4). This consistency allows to read out the compartmental kinetics of each cell population after bleomycin treatment for all substrains combined.

Inflammatory cell kinetics after bleomycin-induced lung injury are robust and reproducible

Analysis of the inflammatory response in the BALF, identified a non-resolving inflammatory response, with the total number of inflammatory cells continuing to increase over the investigated time course of 21 days. In the lung tissue, inflammation was characterized by an immediate increase at day 3, stagnating to day 14 and mostly resolved 21 days post bleomycin exposure (Fig. 4). This suggests that the inflammatory response is persistent, yet compartment dependent.

Early inflammatory changes were mostly dominated by the innate immune system, including both immature and mature neutrophils, monocyte-derived alveolar and interstitial macrophages. In contrast we observed a concomitant decrease in alveolar macrophages. Interestingly, the (inverted) trajectories of alveolar macrophages were comparable to the rise in monocyte-derived macrophages, suggesting a functional replacement by the latter and supports observations in earlier studies (Misharin et al., 2017). Following the rapid increase in the first line responders, neutrophils, their numbers later stagnated or gradually decreased, and even returned to baseline levels in the lung tissue. We also identified a time-dependent increase in SiglecF$^+$ neutrophils following bleomycin application. These cells have recently

been described to be important for cancer progression (Engblom et al., 2017). Similarly, eosinophils and dendritic cells (EOS, DC) exhibited a bell-shape response curve. In contrast monocyte populations (both constitutive and inflammatory) exhibited a slower, but consistent, step wise increase over time, which could be attributed to their contribution to both the innate and adaptive immunity and their role in tissue repair.

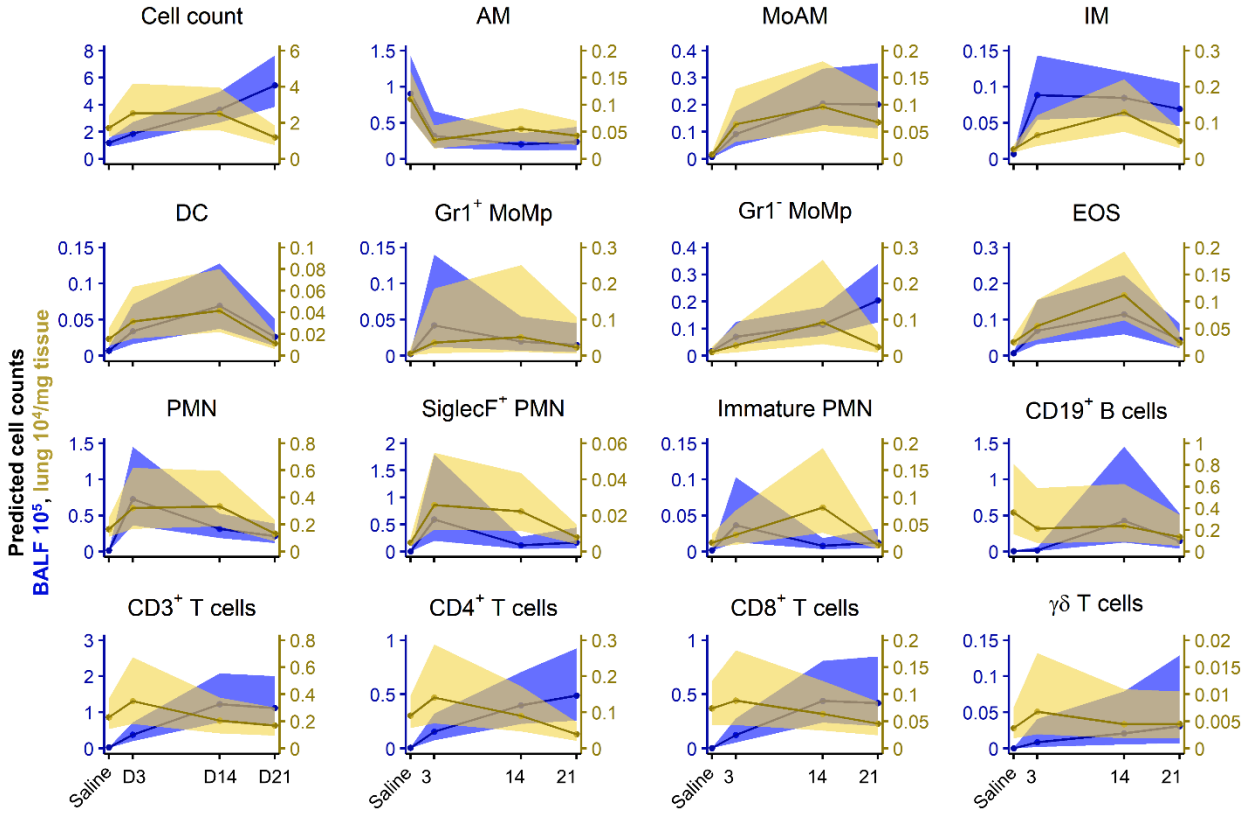

Fig. 4. Non-linear mixed models reveal complex immune cell dynamics occurring in the lung following bleomycin induced lung injury. Plot of back transformed, fitted cell counts and their 95 % confidence intervals using the simplified mixed model [*SalineDay*+*Substrain*~1|*Exp_ID*] of LOG_{x+1} transformed cell counts for BALF (counts·10^5) and lung tissue (counts·10^4/mg tissue). Animal numbers were in BALF in total n = 159 (Saline 60; 3d 23; 14d 39; 21d 37) and in lung in total n = 144 (Saline 56; 3d 23; 14d 32; 21d 33).

At later time points, inflammation was dominated by immune cells from adaptive immunity, with a clear preference to the alveolar compartment. While $CD3^+$ T lymphocytes ($CD4^+$ and $CD8^+$ T cells, respectively) had a steep, yet non-resolving, rise early in the inflammatory response, the $CD19^+$ B cells peaked at 14 days post bleomycin challenge. Interestingly, at the latest time point, 21 days, B cells numbers still continued to rise, implicating their involvement at later stages in this model (Fig. 4A).

Taken together, the multiple inflammatory cell populations show dynamic and distinct inflammatory kinetics with clear compartment preferences. With time the involved immune cells shifted from the innate to the adaptive immune system, with the BAL being more prominently affected then the lung tissue. After 21 days the inflammatory profile was still chronically altered.

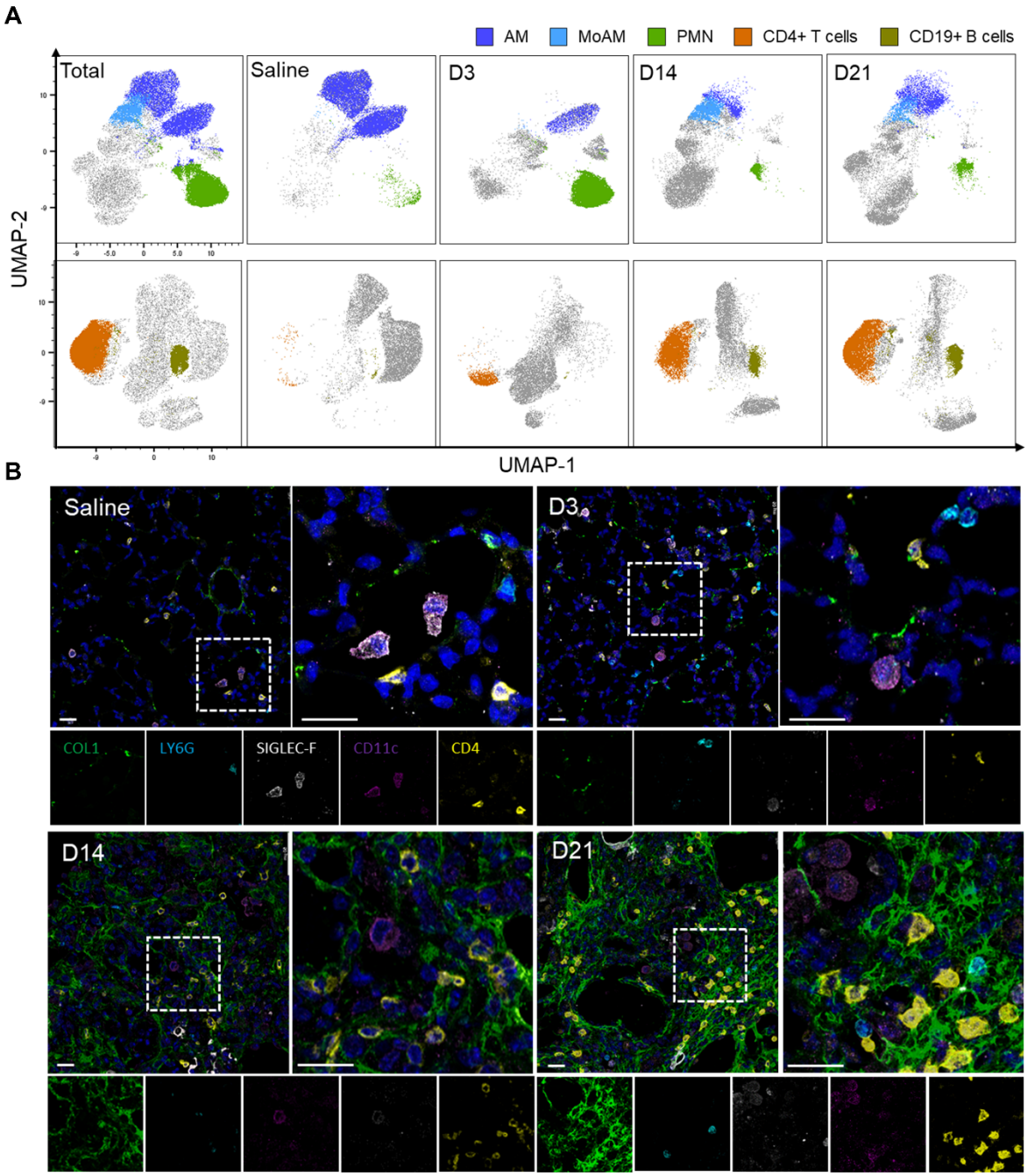

Fig. 5. Temporal and spatial localization of inflammatory cell kinetics in BALF and lung tissue. (A) Uniform Manifold Approximation and Projection (UMAP) plots of concatenated $CD45^+$ populations (min 3 independent samples with max 10'000 $CD45^+$ cells per sample) with overlaid manually gated populations in BALF. (B) Spatial localisation of alveolar macrophages ($CD11c^+$/$SiglecF^+$), neutrophils ($LY6G^+$), and $CD4^+$ T cells during the time course of bleomycin challenge. Nuclei are stained with DAPI (dark blue). Representative pictures of three independent mice at each time point. D3, D14 and D21, represent days 3, 14 and 21 post-bleomycin treatment, respectively.

Based on these results we went back to our FCM data and visualised the kinetics of the most dynamically altered populations via computational FCM (Fig. 4B). As predicted in our modelling data AM populations strongly decreased after bleomycin exposure, while the innate PMNs vastly expanded after 3 days. Adaptive immune cells such as $CD4^+$ T cells, $CD8^+$ T cells and $CD19^+$ B cells expanded

more at later time points and were virtually absent in saline treated mice (Fig. 5A). Visualisation using multi-colour immunofluorescence revealed the co presence of $CD11c^{+/}SiglecF^{+}$ AM, $CD4^{+}$ T cells and $Ly6G^{+}$ neutrophils in fibrotic lung tissue (Fig. 5B), the spatiotemporal presence of these cells point toward close interplay between inflammatory components.

The inflammatory cell landscape continually evolves following bleomycin exposure

The combination of unsupervised multivariate methods and univariate NLME identified the kinetics of each cell type with an early innate response followed later by adaptive immune response. However, the question how the entire landscape differs between different timepoints or which cell types define each stage is still open. In order to answer these questions, we applied three robust machine learning approaches.

Our first approach, OPLS-DA separates the dataset into predictive and non-predictive components. Predictive means the ability to discern between groups in the given classification factor, which was here *SalineDay* {Saline,3,14,21}. The OPLS-DA model quality was thoroughly investigated by cross-validation and permutations tests showing that in both compartments the models were highly significant ($Q2>50$ %, $p<0.001$). Similar to our PCA results (Fig. 3), the inflammatory reaction was more pronounced in the BALF than in the lung, as apparent from a clearer group separation, higher percentages of variability in the predictive component and higher predictive ability (Q2; Fig. 6A). In BALF, the inflammatory landscape at 14 and 21 days post bleomycin were very similar, but very different from the saline controls, while the landscape at 3 days bridged these two poles.

We next investigated conditional inference trees and random forest models to infer which cell populations were the driving factors behind the group differences. Conditional inference trees in the BALF demonstrated that $CD3^{+}$ T cells levels separated early (Saline, D3) and later timepoints (D14, D21). Separating samples on low and high $CD19^{+}$ B cells distinguishes between days 14 and 21, respectively. On the other hand, low levels of PMN strongly predicts saline treated mice and the combination of low AM and $SiglecF^{+}$ PMN aiding the separation between saline, D3 and D14 (Fig. 6B). In the lung compartment, both innate cells (MoAM, AM) and adaptive ($CD4^{+}$ T cells and $CD19^{+}$ B cells) were needed to define the different groups. Saline mice were defined by low levels of MoAM and high AM, while bleomycin treatment by high MoAM and $CD4^{+}$ T cells. Similar to the BALF, day 21 was marked by high $CD19^{+}$ levels, while D14 by was defined by lower B cell and MoAM levels (Fig. 6B). A combination of low MoAM and low AM defined day 3.

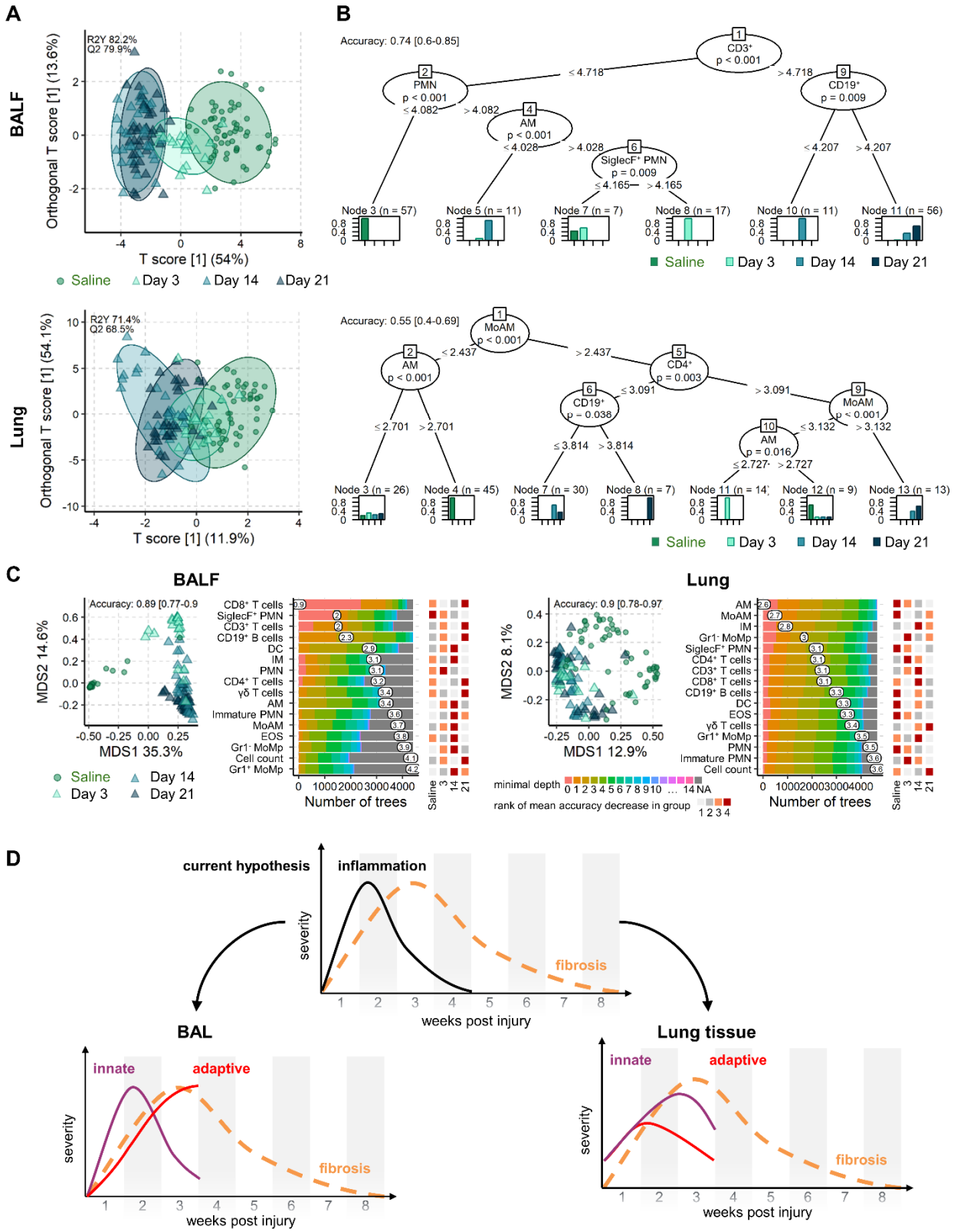

Fig. 6. Exploration of inflammatory cell landscape differences with machine learning in BALF and lung tissue. (A) Scores plot of OPLS-DA models per compartment for the factor *SalineDay* {Saline,3,14,21} with 95 % confidence ellipses for each group. The predictive ability of the models Q^2 was calculated by 7-fold cross validation and 1000 permutation tests reconfirmed model significance with p<0.001. (B) Conditional inference trees per compartment, showing cell types and cut-offs that define each group; saline, days 3, 14 and 21 post bleomycin treatment (*SalineDay*). Model accuracy was evaluated with a stratified split into 65 % trainings and 35 % test set. (C) MDS plot (left panel) of the

proximity matrix of random forest models grown with 5000 trees. Model accuracy was evaluated with a stratified split into 65 % trainings and 35 % test set. The distribution of the minimal depth is shown for each cell type according to the number of trees, the mean of the minimal depth is shown (middle panel). The rank of the mean decrease in accuracy within each group is shown for each cell population (right panel). Animal numbers in all models from A-C were in BALF in total n = 159 (Saline 60; 3d 23; 14d 39; 21d 37) and in lung in total n = 144 (Saline 56; 3d 23; 14d 32; 21d 33). Models were based on LOG_{x+1} transformed cell counts for BALF (counts·10^5) and lung tissue (counts·10^4/mg tissue). (D) Schematic, abstracted summary of all previous finings differentiating between the compartments BAL and lung tissue.

Random forest models were then used to compare the ability of all cell populations to drive group separation. In agreement with previous results, again group separation was clearer in BALF than in lung, as demonstrated by multi-dimensional scaling plots of the random forest proximity matrix and higher accuracy (Fig. 6C). In BALF especially the adaptive immune cells $CD8^+$ and $CD3^+$ T cells as wells as the innate $SiglecF^+$ PMN differed most, as became apparent from their low minimal depth. Between the different groups high $CD8^+$, $CD3^+$ and $CD19^+$ levels were most predictive for late inflammation while low $SiglecF^+$ PMN levels were most predictive for the cellular landscape in saline samples. The random forest suggests some fine but distinct differences between the global inflammatory landscape 14 and 21 days after bleomycin exposure (Fig. 6C). Although both are highly inflamed (OPLS-DA), higher levels of adaptive cells are rather predictive for day 21 (e.g. all T and B cells), while higher levels of some innate cells are more predictive for day 14 than day 21 (e.g. DC, IM, immature PMN, MoAM, EOS) or day 3 (PMN). In contrast, lung models were dominated by macrophage cell populations differing most between the inflammatory stages, foremost the depletion of alveolar macrophages. The random forest models underline that the inflammatory landscape differs notably between lung and BALF.

Discussion

In this study, we have combined computation FCM, advanced data modelling and machine learning approaches to conclusively define the inflammatory cell kinetics following bleomycin treatment in mice. By combining the data from 15 independent experiments, we amassed very large sample numbers, which were far in excess of those normally found in animal experiments. The aggregation of historical samples inherently led to an unbalanced experimental design, which was handled by sophisticated, robust statistical methods. By using pre-processing techniques such as data transformation, we could substantially improve analysis power, which crucially contributed to clearer data interpretation. Changes in the inflammatory profile was dissected using multivariate and univariate statistical methods including non-linear mixed models. Only by applying these techniques in unison were we able to create the most comprehensive picture of inflammatory cell trajectories to date and characterise the sustained inflammation in the bleomycin model of pulmonary fibrosis.

FCM data is normally highly asymmetric i.e. it has many larger values but no values smaller than zero, this non-normal distribution prevents the use of more powerful analysis methods. To re-establish normality we trialled several transformations, but ultimately settled on LOG_{x+1} as it normalised the data distribution, can be easier to interpret and also slightly improved the scedasticity compared to 4RT. Our data modelling approach resulted in a very large sample size, which notably increased statistical power and outweighed the potential drawbacks of added confounding variation from experimental runs or the use of different substrains. Furthermore, when experimental covariance was accounted for as random factor in NLME models, the inflammatory profiles in the BALF and lung tissue of all saline treated animals, irrespective of experiment, were sufficiently similar to be combined into one large control group. Secondly, the trajectories of inflammatory cell profiles were found to be consistent for all five substrains, although their magnitudes slightly differed, which is important for experimental reproducibility in light of using different knockout lines or mice sourced from different companies.

The application of unsupervised and supervised as well as multivariate and univariate, demonstrated how the changes for most populations were more prominent in the BALF than in lung tissue, although the majority of populations showed consistent trajectories in both compartments. In healthy mice, the vast majority of cells in the BALF are alveolar macrophages, while in the lung tissue even at baseline conditions, a highly heterogenous pool of inflammatory cells exists, including macrophages neutrophils, T and B cells. Due to the higher sensitivity of the BALF to monitor inflammatory changes, we would recommend this as the compartment of choice for the majority of cells in FCM analysis. Analysis of the BALF has further advantages such as being easily accessible, without need for additional tissue digestion steps.

Our comprehensive analysis of multiple inflammatory cell population at several time-points, describes the kinetics not only during disease development but also when it is fully established. The initial

inflammatory phase after bleomycin exposure was dominated by early responder cell types from the innate immune system of the myeloid lineage. Neutrophils constitute the first line defence of the immune system and consequently show very acute kinetics, being rapidly recruited and also being the first cell type to resolve, visible as pronounced decreases from day 3 to day 14 after the challenge. In contrast, cells from adaptive immune system, such as B and T cells, increased much slower but continue to expand even at 21 days. The worth of subtyping cell populations is apparent by the inverse kinetics displayed within macrophages, which is only possible by using multicolour analysis. We could show that while the numbers of alveolar macrophages (AlvMp) quickly decrease, monocyte derived macrophages (MoAM) increase. These contrary trajectories would explain the early observation that macrophages numbers were unchanged in this model (Izbicki et al., 2002), but the closer analysis of macrophage subtypes revealed strong dependent changes, as shown by (Misharin et al., 2017, 2013) and now reconfirmed by our results.

Increasingly macrophage heterogeneity has been suggested to play an important role in the pathogenesis of lung fibrosis and have implications for therapeutic strategies. MoAM undergo marked transcriptional changes during their differentiation in the injured lung tissue. These changes are not only associated with a continuous down-regulation of genes typically expressed in monocytes and up-regulation of genes expressed in alveolar macrophages but also with markedly elevated expression of proinflammatory and profibrotic genes related to M1 and M2 phenotype. This unique transcriptomic signature of MoAM provides an explanation how bleomycin-induced lung fibrosis is attenuated following selective depletion of these cells (Joshi et al., 2020; McCubbrey et al., 2018; Misharin et al., 2017). Interestingly, the existence of common profibrotic pathways in MoAM harvested from mice during fibrosis development and profibrotic macrophages obtained from the lungs of IPF patients has been reported (Aran et al., 2019; Misharin et al., 2017). All these observations strongly suggest that selective targeting profibrotic macrophages, rather than the M1 or M2 phenotype, is more likely to be of benefit in such a complex disease as IPF. The potential contribution of MoAM to the resolution of lung fibrosis is still open and remains the subject of future studies. Recent data supports this hypothesis (Cui et al., 2020). Hence, MoAM could represent very plastic cell population with distinct functions in different phases of lung fibrogenesis.

Early and late fibrotic stages were characterized by increased numbers of T and B cells in the BALF, while numbers in the lung tissue remained relatively stable, this reflects earlier reports describing the presence of T cells in IPF lungs (Balestro et al., 2016; Todd et al., 2013). Here B cells are of particular interest, as abnormal B cell aggregates have been described in IPF lungs (Marchal-Sommé et al., 2006) and diverse circulating IgG autoantibodies were found in IPF plasma (Kurosu et al., 2008; Ogushi et al., 2001; Taillé et al., 2011). Furthermore, individual auto-immunoglobulins were linked to severity and/or poor prognosis of IPF (Kahloon et al., 2013; Ogushi et al., 2001) thus suggesting the causal role of certain autoantigens in IPF. Accordingly, transcriptome-profiling of lung tissue derived from

pirfenidone-treated patients revealed downregulation of B cell related genes (Kwapiszewska et al., 2018). Future studies will, however, demonstrate whether these findings open an exciting new avenue for immunotherapy-based approaches in IPF.

Limitations

Despite analysing three independent timepoints, which cover the major stages of the bleomycin model, some timepoints are still missing. However, we consciously wanted to reuse existing experiments and avoid sacrifice of new animals. Future investigation would profit from an expansion, e.g. by inclusion of existing measurements from other groups, to cover also the progression from the initial inflammation towards active fibrosis phase by including analysis at day seven. Similarly, inflammatory profiling during fibrosis resolution, i.e. after 28 or 35 days, would deliver valuable insights on the involvement of specific subtypes during resolution. From a statistical point of view, the unbalanced study design with differing sample numbers in subgroups is unfavourable, which complicates analysis and loses some power. However, our use of robust methods such as NLME and machine learning methods (random forest) were able to overcome these limitations. Although over a dozen independent experimental runs were included, this is not a multi-centric study. Quantitative comparison of results from other laboratories at other sites and other strains/substrains would allow to even better explore bleomycin model system robustness and reproducibility. In this study, manual gating was used to identify different cell populations, thereby including expert knowledge into the analysis and gating specificity was confirmed shown by UMAP overlays. For some populations in the UMAP plots (e.g. AM), the populations were more spread than expected, this was most likely due to do different marker intensity (in this case CD11c) between different experimental runs. The topic of auto-gating is rapidly developing and promises to considerably save hands-on time and foremost the potential to detect rare, otherwise undetected cell subpopulations. The focus of this study was to primarily determine the inflammatory cell kinetics, however to further unravel the role of inflammation and potential therapeutic targets in fibrosis a quantified link of cell subpopulations to fibrotic processes is warranted.

Recommendations

This study explored fundamental aspects of the bleomycin animal model with good power owing to the high sample numbers so that constructive recommendations can be inferred.

(I) In order to ascertain technical success of the experiment we strongly recommend to always include a negative control (saline) and a positive control (bleomycin, transgenic or knock out otherwise untreated) group with each $n \geq 8$. Foremost this serves to rate the strength of induced fibrosis and technical quality of the experiment. Statistical power gain is very high for every added sample in the

single digit region. An n# of ≥8 leaves some safety margin to stay above the critical level of n=5 to handle the occasional, unavoidable loss of samples due to premature death or technical problems.

(II) For more sensitive and pronounced inflammatory readouts BALF should be routinely sampled together with lung tissue and both samples should be subjected to analogous analysis.

(III) For subsequent statistical analysis we strongly recommend to investigate distribution and potential for improvements from data transformations and especially for multivariate methods additional centring and scaling. We also recommend to use both multivariate and univariate, unsupervised and supervised methods as they complement each other well in their type of generated insights.

Conclusions

The measurement of inflammatory cellular landscapes in the bleomycin-induced lung-injury mouse model with flow cytometry is very robust and suitable to quantify kinetic changes in multiple cell populations simultaneously. The results allowed to infer recommendations such as to add negative and positive control, apply data pre-processing, combine multivariate and univariate methods and to routinely also investigate BALF. We also found that the unintended development of potential substrains does not per se hinder general reproducibility of results and the approach to adapt bleomycin doses to the current experimental run is viable. This study underlines the relevance of combined analysis for more holistic insights into inflammatory profile changes. Cell populations show quite distinct trajectories in their kinetics. We also conclude that inflammatory cell-based response is active before, during and after manifestation of fibrosis with a shift from the initial innate immune cell domination towards the adaptive arm and inflammatory cell accumulation is not resolved after 3 weeks.

Material and Methods

Animals

All animal experiments were approved by the local authorities (Austrian Ministry of Science, Research and Economics) (BMWF-66.010-0038-II-3b-2013, BMWFW- 66.010/0038-WF/II/3b/2014, BMWFW-66.010/0049-WF/V/3b/2017, 66.010/0177-WF/3b/2017) and were performed in accordance with relevant guidelines and regulations. Wild type groups of several different, independent experiments (unpublished and published (Biasin et al., 2017)) were pooled and analysed. For each experimental run wild type mice were obtained from Charles River or bred in-house in case of wild type littermates and are annotated as separate strains. Overview of all strains and group sizes is given in Supplementary Table S1. All mice were maintained with 12 h light/ dark cycles and they had access to water and standard chow *ad libitum*.

Bleomycin challenge and animal handling

Male mice (25-30 g body weight) were anesthetized with isoflurane 2–2.5 % and intra-tracheal administered with bleomycin (Sigma, Vienna, Austria) or saline solution (0.9 % w/v NaCl) using a MicroSprayer® Aerosoliser (Penn-Century Inc., PA, Pennsylvania, USA), as previously described (Biasin et al., 2020, 2017). Each bleomycin lot was titrated to give a comparable response for each strain; dose range was 0.7-3.5 U/kg b.w., Supplementary Data 1). After bleomycin instillation, mice were closely monitored till they completely recovered from anaesthesia. Bleomycin or saline solution administration was performed once and animals were sacrificed after 3, 14 or 21 days.

BALF and lung tissue preparation for flow cytometry

Mice were euthanized via exsanguination and the lungs were perfused with phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.2), through the right ventricle. Mice were then lavaged with 1 ml PBS containing the Pierce protease inhibitor cocktail (ThermoFisher Scientific, Vienna, Austria) and 1 mM EDTA. The obtained BALF was centrifuged, washed with 1 ml MACS buffer (2 mM EDTA, 0.5 % BSA in 1X PBS), before being resuspended in 0.5 ml for cell counting and consequent FCM staining. Single cell lung tissue homogenates were performed as previously described (Nagaraj et al., 2017). Briefly, the lower right lobe was weighed, cut into approximately 1 mm pieces and digested with 0.7 mg/ml Collagenase and 30 µg/ml DNAse in RPMI medium supplemented with 10 % FCS, 2 mM glutamine and 1 % penicillin-streptomycin (ThermoFisher Scientific) for 40 min at 37 °C with rotation at 350 rpm. The minced tissue was passed through a 100 µm cell strainer to obtain a single cell suspension. In case of red blood cells contamination, the cell suspension was treated with erythrolysis buffer (2.6 mM NH_4Cl, 0.09 M KCO_3, 0.6 M EDTA) for 5 min at room temperature. The number of live cells were counted using trypan blue

exclusion and then stained with fixable viability stain (ThermoFisher Scientific), washed and then fixed with 1 % paraformaldehyde for 15 min on ice before being resuspended in MACS buffer.

Flow cytometry

Single cell suspensions were initially incubated with an Fc-receptor-binding antibody (ThermoFisher Scientific) for 5 min on ice to prevent nonspecific binding. A master-mix containing one of two different antibody combinations against cell surface markers (Supplementary Table S2) was added to the cells incubated for 20 min at 4 °C. For each sample between 30'000 and 300'000 events were recorded on a LSRII Flow Cytometer (BD Biosciences, Vienna, Austria) or Cytoflex S (Beckman Coulter, Vienna, Austria). Samples were analysed either using FACSDiva (BD Biosciences) or FlowJo v10.6.2 (LLC, Ashland, Oregon) software by users blinded to treatment condition. Cells were initially gated on FSC and SSC characteristics and duplexes were removed using FSC-A / FSC-H dot blot, dead cells were gated out using viability exclusion. Cells positive for the pan-leukocyte marker CD45 were taken for further analysis, cell populations were identified using the gating strategy (Fig. 1C and Table 1), as described in the results and based on published studies (Biasin et al., 2017; Gungl et al., 2018; Misharin et al., 2017, 2013; Tighe et al., 2019b). A complete description of all antibodies is given in Supplementary Table S2. Cell numbers are reported 10^5 in the BALF and 10^4/mg tissue for the lung. Uniform Manifold Approximation and Projection (UMAP) plots were performed in FlowJo, using default settings (nearest neighbours 15, minimum distance value 0.5, Euclidean distance). First, fcs files from at least three individual mice per analysis timepoint were downsampled to max 10'000 events and then concatenated. Manually gated populations were then overlaid on UMAP plots to determine they kinetics.

Trichrome and immunofluorescence staining

After BALF, the lungs were inflated with 4 % formalin via the trachea and then paraffin embedded. Slides were cut at 2.5 µm thick and stained with Masson's trichrome according to standard protocols. Slides were scanned and imaged with a Virtual Slides VS120 Microscope and OlyVia Software (both from Olympus, Vienna, Austria). For multi-colour immunofluorescence staining, 2.5 µm paraffin-embedded lung sections were dewaxed and subjected to heat induced antigen retrieval at pH6 (Perkin-Elmer, Waltham, MA) using an antigen retrieval chamber for 15 min at 200 W. Slides were blocked with Perkin-Elmer Antibody Block solution for 20 min in a humidified chamber, and primary antibodies (Supplementary Table S3) were sequentially incubated o/n 4 °C in Perkin-Elmer Antibody Diluent. After washing with TBS-T (274 mM NaCl, 47.6 mM Tris HCl + 2 % v/v Tween20 in H_2O) primary antibodies against CD4, SiglecF and CD45 were detected with the Opal Polymer HRP secondary antibody (Perkin-Elmer), using the Opal 540, 620, 690 substrates, respectively. Antibodies against Collagen I, CD11c and CD45 were used simultaneously and detected with AlexaFluor-conjugated secondary antibodies, donkey anti-goat AlexaFluor488, donkey anti-rabbit AlexaFluor555, chicken

anti-rat AlexaFluor647, respectively. Nuclear counterstaining was performed with DAPI solution 1 mg/ml (ThermoFisher Scientific).

Confocal imaging

For imaging immunofluorescence stained slides, a Leica TCS-SP8 (DMi8 inverted microscope with a LIAchroic scan head) lightning confocal microscope was used (Leica, Wetzlar, Germany). The acquisition process followed a "sequential workflow" with well-defined settings (shown in Supplementary Table S4). In order to minimize fluorescent overlap the plugin "Channel Dye Separation" of Leica Imaging system was used. The following objectives were used: Plan Fluotar 20x/0.75 multi immersion objective and Plan Fluotar 40x/1.25 glycerol immersion objective. Images were acquired at 2048 x 2048 and a pixel size of 142 x 142nm.

Statistical analysis

Data visualisation and statistical analysis were performed with R v3.6.3 (R Core Team, 2020) (using the packages readxl, openxlsx, plyr, stringr, tidyr, reshape, colorspace, RColorBrewer, ggplot2, ggpubr, ggrepel, gridExtra, magrittr, cowplot, plotly, lemon, lawstat, dendsort, pheatmap, cellWise, missMDA, FactoMineR, nlme, emmeans, MetaboAnalystR 2.0, caret, randomForest, randomForestExplainer, partykit, e1071), TIBCO Spotfire v10.9.0, TIBCO, Palo Alto, CA and FlowJo v10 (LLC, Ashland, Oregon).

All reported p-values were adjusted for multiple testing according to Benjamini-Hochberg (BH) denoted as p_{BH} (R function *p.adjust*). Distribution and scedasticity were investigated with Kolmogorov-Smirnov test and Brown-Forsythe Levene-type test, respectively (p_{BH} Supplementary Data 1). Seven common transformations were tested: square root, reciprocal, Freeman Tukey, logit (on counts mapped to 0.25-0.75), LOG, LOG_{x+1}, 4RT (Supplementary Fig. S1).

Principal component analysis (PCA) analysis (R function *prcomp*) was performed centred and scaled to unit variance (z-scaled) on total cell counts (untransformed, LOG_{x+1} or 4RT transformed). The dataset (303 samples, 16 cell populations) contained no missing values and 1.3 % zeros. MacroPCA analysis (R function *MacroPCA*) was performed centred and scaled to unit variance on total cell counts (untransformed, LOG_{x+1} or 4RT transformed). The number of components was set to cumulatively retain 80 % of explained variance, but to deliver between two and ten components. Hierarchical clustering analysis was performed centred and scaled to unit variance (R function *scale*) on total cell counts, for untransformed data per cell type than samples. LOG_{x+1} or 4RT data was centred and scaled only per cell type. The dendrograms were clustered by Lance-Williams dissimilarity update with complete linkage (R function *dist* and *hclust*) and sorted (R function *dendsort*) at every merging point according to the average distance of subtrees and plotted at the corresponding heat maps (R function *pheatmap*).

548 Non-linear mixed models were fitted (R function simple models *gls* or mixed models *lme* with
549 maximum likelihood (ML), with LOGx+1 transformation and no longitudinal covariance applied (mice
550 were sacrificed at each time point). Model selection was based on the forward addition approach and
551 complex models were rechecked by backward dropping of factors. Simple models were constructed
552 using the forward addition approach incorporating the fixed factors *Treatment* {Saline,Bleo}, *Day*
553 {3,14,21} post treatment and the mouse background, *Substrain* {A,B,C,D,E}. The interactions,
554 *Treatment:Substrain and Treatment:Day* were include to determine whether the treatment effect
555 depended on the *Substrain* or *Day*. Mixed models additionally included the experimental ID as a random
556 factor (~1|Exp_ID). Complex mixed models were created by combining mixed models with the
557 interactions *Treatment:Substrain* and/or *Treatment:Day*. Models were then simplified by merging all
558 saline samples into one control group generating the simple model [*SalineDay+Substrain*] and by
559 including *Exp_ID* as a random factor the mixed model [*SalineDay+Substrain~1| Exp_ID*]. Due to rank
560 deficiencies arising from the unbalanced design the model *SalineDay:Substrain* was not possible.
561 Criteria for model performance and suitability were lower AIC (Akaike information criterion; relative
562 estimate of information loss), higher log-likelihood (goodness of fit), significance in log likelihood ratio
563 test comparing two models, quality of Q-Q plots and randomness in residual[Box 2] plots (Supplementary
564 Data 1 and Supplementary Fig. S2). Post-hoc pairwise comparisons were readout as back transformed
565 estimates (R function *emmeans*, type = "response") with $p_{BH} \leq 0.05$ being considered statistically
566 significant.

567 Orthogonal projections to latent structures discriminant analysis (OPLS-DA) on LOG_{x+1} data was
568 performed centred and scaled to unit variance (R function *Normalization* with scaleNorm="AutoNorm"
569 and R function *OPLSR.Anal*) with a standard 7-fold cross validation for the classification factor
570 *SalineDay*. Model stability was additionally verified with 1000 random label permutations.

571 Conditional inference trees were fit with default settings (R function *ctree*) which limits tree size to
572 include only significant splits avoiding overfitting, so that no further cross-validation or pruning was
573 applied. The random forest (R function *randomForest*) error rates decrease markedly within the first
574 100 trees and stabilized fully after 1500 to 2500 trees. All reported random forests grown with 5000
575 trees to guarantee stability and hyperparameter, mtry (8 in BALF and 2 in lung) was tuned to minimal
576 out-of-bag errors (OOB) (R function *tuneRF*). The model stability and prediction quality (R function
577 *confusionMatrix*) of conditional inference trees and random forest was evaluated by splitting the
578 LOG_{x+1} randomly into trainings/test set (65 % / 35 %) stratified for the classification factor *SalineDay*
579 (R function *createDataPartition*).

580

Author contributions

Conceptualisation, Data curation, Formal analysis, Software, Visualisation and Validation, N.B and L.M.M; Methodology, L.M.M; Investigation, V.B, B.M.N, F.V, N.S, K.J and L.M.M; Resources, V.B, B.M.N., N.S., K.J., G.K. and L.M.M.; Writing – original draft, N.B, V.B, G.K, M.W, L.M.M; Writing – Review & Editing, all authors; Project administration, Supervision and Funding acquisition, G.K. and L.M.M.

Acknowledgements

We are very grateful for the technical assistance from Kerstin Schweighofer, Sabine Halsegger, Thomas Fuchs, Nina Treitler, Eva Grasmann, Camilla Götz, and Betty Schrenk, as well as Slaven Crnkovic, Mathias Hochgerner, Chandran Nagaraj and Horst Olschewski for the valuable discussions and advice. Diana Schnögl, Francesco Valzano and Neha Sharma are members of the MolMed graduate programme. Diana Schnögl is funded by FFG, project number 870904 awarded to Leigh Marsh. Francesco Valzano is funded by FFG, project number 874229 awarded to Grazyna Kwapiszewska. Valentina Biasin is funded by Austrian Science Fund (FWF; project number, T1032-B34).

Competing interests

The authors declare that they have no competing interests.

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the paper or the Supplementary Materials.

References

Aran D, Looney AP, Liu L, Wu E, Fong V, Hsu A, Chak S, Naikawadi RP, Wolters PJ, Abate AR, Butte AJ, Bhattacharya M. 2019. Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* **20**:163–172. doi:10.1038/s41590-018-0276-y

B Moore B, Lawson WE, Oury TD, Sisson TH, Raghavendran K, Hogaboam CM. 2013. Animal models of fibrotic lung disease. *Am J Respir Cell Mol Biol* **49**:167–79. doi:10.1165/rcmb.2013-0094TR

Balestro E, Calabrese F, Turato G, Lunardi F, Bazzan E, Marulli G, Biondini D, Rossi E, Sanduzzi A, Rea F, Rigobello C, Gregori D, Baraldo S, Spagnolo P, Cosio MG, Saetta M. 2016. Immune Inflammation and Disease Progression in Idiopathic Pulmonary Fibrosis. *PLoS One* **11**:e0154516. doi:10.1371/journal.pone.0154516

Biasin V, Crnkovic S, Sahu-Osen A, Birnhuber A, El Agha E, Sinn K, Klepetko W, Olschewski A, Bellusci S, Marsh LM, Kwapiszewska G. 2020. PDGFRα and αSMA mark two distinct mesenchymal cell populations involved in parenchymal and vascular remodeling in pulmonary fibrosis. *Am J Physiol Cell Mol Physiol* ajplung.00128.2019. doi:10.1152/ajplung.00128.2019

Biasin V, Wygrecka M, Marsh LM, Becker-Pauly C, Brcic L, Ghanim B, Klepetko W, Olschewski A, Kwapiszewska G. 2017. Meprin β contributes to collagen deposition in lung fibrosis. *Sci Rep* **7**:39969. doi:10.1038/srep39969

Cui H, Jiang D, Banerjee S, Xie N, Kulkarni T, Liu R-M, Duncan SR, Liu G. 2020. Monocyte-derived alveolar macrophage Apolipoprotein E participates in pulmonary fibrosis resolution. *JCI Insight*. doi:10.1172/jci.insight.134539

Della Latta V, Cecchettini A, Del Ry S, Morales MA. 2015. Bleomycin in the setting of lung fibrosis induction: From biological mechanisms to counteractions. *Pharmacol Res* **97**:122–130. doi:10.1016/j.phrs.2015.04.012

El Agha E, Kramann R, Schneider RK, Li X, Seeger W, Humphreys BD, Bellusci S. 2017. Mesenchymal Stem Cells in Fibrotic Disease. *Cell Stem Cell*. doi:10.1016/j.stem.2017.07.011

Engblom C, Pfirschke C, Zilionis R, Da Silva Martins J, Bos SA, Courties G, Rickelt S, Severe N, Baryawno N, Faget J, Savova V, Zemmour D, Kline J, Siwicki M, Garris C, Pucci F, Liao HW, Lin YJ, Newton A, Yaghi OK, Iwamoto Y, Tricot B, Wojtkiewicz GR, Nahrendorf M, Cortez-Retamozo V, Meylan E, Hynes RO, Demay M, Klein A, Bredella MA, Scadden DT, Weissleder R, Pittet MJ. 2017. Osteoblasts remotely supply lung tumors with cancer-promoting SiglecFhigh neutrophils. *Science (80-)* **358**. doi:10.1126/science.aal5081

Gungl A, Biasin V, Wilhelm J, Olschewski A, Kwapiszewska G, Marsh LM. 2018. Fra2 Overexpression in Mice Leads to Non-allergic Asthma Development in an IL-13 Dependent Manner. *Front Immunol* **9**:2018. doi:10.3389/fimmu.2018.02018

Hubert M, Rousseeuw PJ, Van den Bossche W. 2019. MacroPCA: An All-in-One PCA Method

638 Allowing for Missing Values as Well as Cellwise and Rowwise Outliers. *Technometrics* 1–18.
639 doi:10.1080/00401706.2018.1562989
640 Izbicki G, Segel MJ, Christensen TG, Conner MW, Breuer R. 2002. Time course of bleomycin-induced
641 lung fibrosis. *Int J Exp Pathol* **83**:111–9. doi:10.1046/j.1365-2613.2002.00220.x
642 Joshi N, Watanabe S, Verma R, Jablonski RP, Chen CI, Cheresh P, Markov NS, Reyfman PA,
643 McQuattie-Pimentel AC, Sichizya L, Lu Z, Piseaux-Aillon R, Kirchenbuechler D, Flozak AS,
644 Gottardi CJ, Cuda CM, Perlman H, Jain M, Kamp DW, Budinger GRS, Misharin A V. 2020. A
645 spatially restricted fibrotic niche in pulmonary fibrosis is sustained by M-CSF/M-CSFR signalling
646 in monocyte-derived alveolar macrophages. *Eur Respir J* **55**. doi:10.1183/13993003.00646-2019
647 Kahloon RA, Xue J, Bhargava A, Csizmadia E, Otterbein L, Kass DJ, Bon J, Soejima M, Levesque
648 MC, Lindell KO, Gibson KF, Kaminski N, Banga G, Oddis C V., Pilewski JM, Sciurba FC,
649 Donahoe M, Zhang Y, Duncan SR. 2013. Patients with Idiopathic Pulmonary Fibrosis with
650 Antibodies to Heat Shock Protein 70 Have Poor Prognoses. *Am J Respir Crit Care Med* **187**:768–
651 775. doi:10.1164/rccm.201203-0506OC
652 Kurosu K, Takiguchi Y, Okada O, Yumoto N, Sakao S, Tada Y, Kasahara Y, Tanabe N, Tatsumi K,
653 Weiden M, Rom WN, Kuriyama T. 2008. Identification of Annexin 1 as a Novel Autoantigen in
654 Acute Exacerbation of Idiopathic Pulmonary Fibrosis. *J Immunol* **181**:756–767.
655 doi:10.4049/jimmunol.181.1.756
656 Kwapiszewska G, Gungl A, Wilhelm J, Marsh LM, Puthenparampil HT, Sinn K, Didiasova M,
657 Klepetko W, Kosanovic D, Schermuly RT, Wujak L, Weiss B, Schaefer L, Schneider M, Kreuter
658 M, Olschewski1 A, Seeger W, Olschewski H, Wygrecka M. 2018. Transcriptome profiling reveals
659 the complexity of pirfenidone effects in idiopathic pulmonary fibrosis. *Eur Respir J* **52**.
660 doi:10.1183/13993003.00564-2018
661 Maher TM, Strek ME. 2019. Antifibrotic therapy for idiopathic pulmonary fibrosis: time to treat. *Respir*
662 *Res* **20**:205. doi:10.1186/s12931-019-1161-4
663 Marchal-Sommé J, Uzunhan Y, Marchand-Adam S, Valeyre D, Soumelis V, Crestani B, Soler P. 2006.
664 Cutting Edge: Nonproliferating Mature Immune Cells Form a Novel Type of Organized
665 Lymphoid Structure in Idiopathic Pulmonary Fibrosis. *J Immunol* **176**:5735–5739.
666 doi:10.4049/jimmunol.176.10.5735
667 Marsh LM, Jandl K, Grünig G, Foris V, Bashir M, Ghanim B, Klepetko W, Olschewski H, Olschewski
668 A, Kwapiszewska G. 2018. The inflammatory cell landscape in the lungs of patients with
669 idiopathic pulmonary arterial hypertension. *Eur Respir J* **51**:1701214.
670 doi:10.1183/13993003.01214-2017
671 Marshall DC, Salciccioli JD, Shea BS, Akuthota P. 2018. Trends in mortality from idiopathic
672 pulmonary fibrosis in the European Union: an observational study of the WHO mortality database
673 from 2001–2013. *Eur Respir J* **51**:1701603. doi:10.1183/13993003.01603-2017
674 McCubbrey AL, Barthel L, Mohning MP, Redente EF, Mould KJ, Thomas SM, Leach SM, Danhorn T,

675 Gibbings SL, Jakubzick C V., Henson PM, Janssen WJ. 2018. Deletion of c-FLIP from CD11b ^{hi}
676 Macrophages Prevents Development of Bleomycin-induced Lung Fibrosis. *Am J Respir Cell Mol*
677 *Biol* **58**:66–78. doi:10.1165/rcmb.2017-0154OC
678 Meltzer EB, Noble PW. 2008. Idiopathic pulmonary fibrosis. *Orphanet J Rare Dis* **3**:8.
679 doi:10.1186/1750-1172-3-8
680 Misharin A V., Morales-Nebreda L, Mutlu GM, Budinger GRS, Perlman H. 2013. Flow Cytometric
681 Analysis of Macrophages and Dendritic Cell Subsets in the Mouse Lung. *Am J Respir Cell Mol*
682 *Biol* **49**:503–510. doi:10.1165/rcmb.2013-0086MA
683 Misharin A V., Morales-Nebreda L, Reyfman PA, Cuda CM, Walter JM, McQuattie-Pimentel AC,
684 Chen C-I, Anekalla KR, Joshi N, Williams KJN, Abdala-Valencia H, Yacoub TJ, Chi M, Chiu S,
685 Gonzalez-Gonzalez FJ, Gates K, Lam AP, Nicholson TT, Homan PJ, Soberanes S, Dominguez S,
686 Morgan VK, Saber R, Shaffer A, Hinchcliff M, Marshall SA, Bharat A, Berdnikovs S, Bhorade
687 SM, Bartom ET, Morimoto RI, Balch WE, Sznajder JI, Chandel NS, Mutlu GM, Jain M, Gottardi
688 CJ, Singer BD, Ridge KM, Bagheri N, Shilatifard A, Budinger GRS, Perlman H. 2017. Monocyte-
689 derived alveolar macrophages drive lung fibrosis and persist in the lung over the life span. *J Exp*
690 *Med* **214**:2387–2404. doi:10.1084/jem.20162152
691 Nagaraj C, Haitchi HM, Heinemann A, Howarth PH, Olschewski A, Marsh LM. 2017. Increased
692 Expression of p22phox Mediates Airway Hyperresponsiveness in an Experimental Model of
693 Asthma. *Antioxid Redox Signal* **27**:1460–1472. doi:10.1089/ars.2016.6863
694 Ogushi F, Tani K, Endo T, Tada H, Kawano T, Asano T, Huang L, Ohmoto Y, Muraguchi M, Moriguchi
695 H, Sone S. 2001. Autoantibodies to IL-1 alpha in sera from rapidly progressive idiopathic
696 pulmonary fibrosis. *J Med Invest* **48**:181–9.
697 Peng R, Sridhar S, Tyagi G, Phillips JE, Garrido R, Harris P, Burns L, Renteria L, Woods J, Chen L,
698 Allard J, Ravindran P, Bitter H, Liang Z, Hogaboam CM, Kitson C, Budd DC, Fine JS, Bauer
699 CMT, Stevenson CS. 2013. Bleomycin Induces Molecular Changes Directly Relevant to
700 Idiopathic Pulmonary Fibrosis: A Model for "Active" Disease. *PLoS One* **8**:e59348.
701 doi:10.1371/journal.pone.0059348
702 R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for
703 Statistical Computing.
704 Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, Fernandez R,
705 Akbarpour M, Chen C-I, Ren Z, Verma R, Abdala-Valencia H, Nam K, Chi M, Han S, Gonzalez-
706 Gonzalez FJ, Soberanes S, Watanabe S, Williams KJN, Flozak AS, Nicholson TT, Morgan VK,
707 Winter DR, Hinchcliff M, Hrusch CL, Guzy RD, Bonham CA, Sperling AI, Bag R, Hamanaka
708 RB, Mutlu GM, Yeldandi A V., Marshall SA, Shilatifard A, Amaral LAN, Perlman H, Sznajder
709 JI, Argento AC, Gillespie CT, Dematte J, Jain M, Singer BD, Ridge KM, Lam AP, Bharat A,
710 Bhorade SM, Gottardi CJ, Budinger GRS, Misharin A V. 2019. Single-Cell Transcriptomic
711 Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am J*

Respir Crit Care Med **199**:1517–1536. doi:10.1164/rccm.201712-2410OC

Rousseeuw PJ, Hubert M. 2018. Anomaly detection by robust statistics. *Wiley Interdiscip Rev Data Min Knowl Discov* **8**:e1236. doi:10.1002/widm.1236

Saeys Y, Van Gassen S, Lambrecht BN. 2016. Computational flow cytometry: helping to make sense of high-dimensional immunology data. *Nat Rev Immunol* **16**:449–462. doi:10.1038/nri.2016.56

Selman M, King TE, Pardo A. 2001. Idiopathic Pulmonary Fibrosis: Prevailing and Evolving Hypotheses about Its Pathogenesis and Implications for Therapy. *Ann Intern Med* **134**:136. doi:10.7326/0003-4819-134-2-200101160-00015

Selman M, Pardo A. 2002. Idiopathic pulmonary fibrosis: an epithelial/fibroblastic cross-talk disorder. *Respir Res* **3**:3. doi:10.1186/RR175

Taillé C, Grootenboer-Mignot S, Boursier C, Michel L, Debray M-P, Fagart J, Barrientos L, Mailleux A, Cigna N, Tubach F, Marchal-Sommé J, Soler P, Chollet-Martin S, Crestani B. 2011. Identification of Periplakin as a New Target for Autoreactivity in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* **183**:759–766. doi:10.1164/rccm.201001-0076OC

Tashiro J, Rubio GA, Limper AH, Williams K, Elliot SJ, Ninou I, Aidinis V, Tzouvelekis A, Glassberg MK. 2017. Exploring animal models that resemble idiopathic pulmonary fibrosis. *Front Med*. doi:10.3389/fmed.2017.00118

Tighe RM, Misharin A V., Jakubzick C V., Brinkman R, Curtis JL, Duggan R, Freeman CM, Herold S, Janssen W, Nakano H, Redente EF, Singer BD, Sperling AI, Swaminathan S, Yu YR, Zacharias WJ. 2019a. Improving the quality and reproducibility of flow cytometry in the lung. *Am J Respir Cell Mol Biol* **61**:150–161. doi:10.1165/rcmb.2019-0191ST

Tighe RM, Misharin A V., Jakubzick C V., Brinkman R, Curtis JL, Duggan R, Freeman CM, Herold S, Janssen W, Nakano H, Redente EF, Singer BD, Sperling AI, Swaminathan S, Yu YR, Zacharias WJ. 2019b. Improving the quality and reproducibility of flow cytometry in the lung. *Am J Respir Cell Mol Biol* **61**:150–161. doi:10.1165/rcmb.2019-0191ST

Todd NW, Scheraga RG, Galvin JR, Iacono AT, Britt EJ, Luzina IG, Burke AP, Atamas SP. 2013. Lymphocyte aggregates persist and accumulate in the lungs of patients with idiopathic pulmonary fibrosis. *J Inflamm Res* **6**:63–70. doi:10.2147/JIR.S40673

Wuyts WA, Agostini C, Antoniou KM, Bouros D, Chambers RC, Cottin V, Egan JJ, Lambrecht BN, Lories R, Parfrey H, Prasse A, Robalo-Cordeiro C, Verbeken E, Verschakelen JA, Wells AU, Verleden GM. 2013. The pathogenesis of pulmonary fibrosis: a moving target. *Eur Respir J* **41**:1207–18. doi:10.1183/09031936.00073012

Xie T, Wang Y, Deng N, Huang G, Taghavifar F, Geng Y, Liu N, Kulur V, Yao C, Chen P, Liu Z, Stripp B, Tang J, Liang J, Noble PW, Jiang D. 2018. Single-Cell Deconvolution of Fibroblast Heterogeneity in Mouse Pulmonary Fibrosis. *Cell Rep* **22**:3625–3640. doi:10.1016/j.celrep.2018.03.010

Supplementary Figures and Tables

A

transformation	equation
square root	$\sqrt[2]{x}$
reciprocal	$\frac{1}{x}$
Freeman Tukey	$\sqrt[2]{x} + \sqrt[2]{x+1}$
logit	$\ln\left(\frac{x_{map}}{(1-x_{map})}\right)$ with $x_{map} = \frac{(x-Y_{lwr})*((upr-d)-(lwr-d))}{(Y_{upr}-Y_{lwr})} + d$
	BALF $Y_{upr} = 30 \cdot 10^5$, lung $Y_{upr} = 15 \cdot 10^4$, $Y_{lwr} = 0$, $upr = 1$, $lwr = 0$, $d = 0.25$
LOG	$log_{10}(x)$
LOGx+1	$log_{10}(x+1)$
4RT	$\sqrt[4]{x}$

x – cell counts (BALF 10^5, lung 10^4/mg tissue)

B

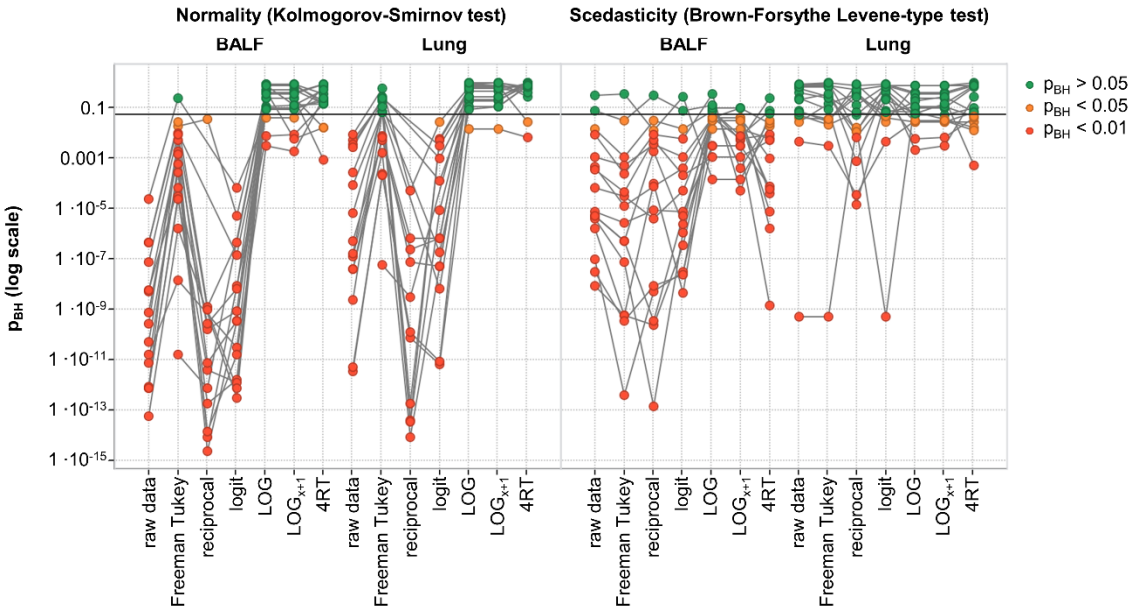

Fig. S1. Data transformations improve distribution and scedasticity. (A) List of tested data transformations with equations. (B) Normality and scedasticity was tested for each of the 16 cell populations in either 159 BALF or 144 lung samples for each of the transformations. The horizontal line denotes pBH= 0.05.

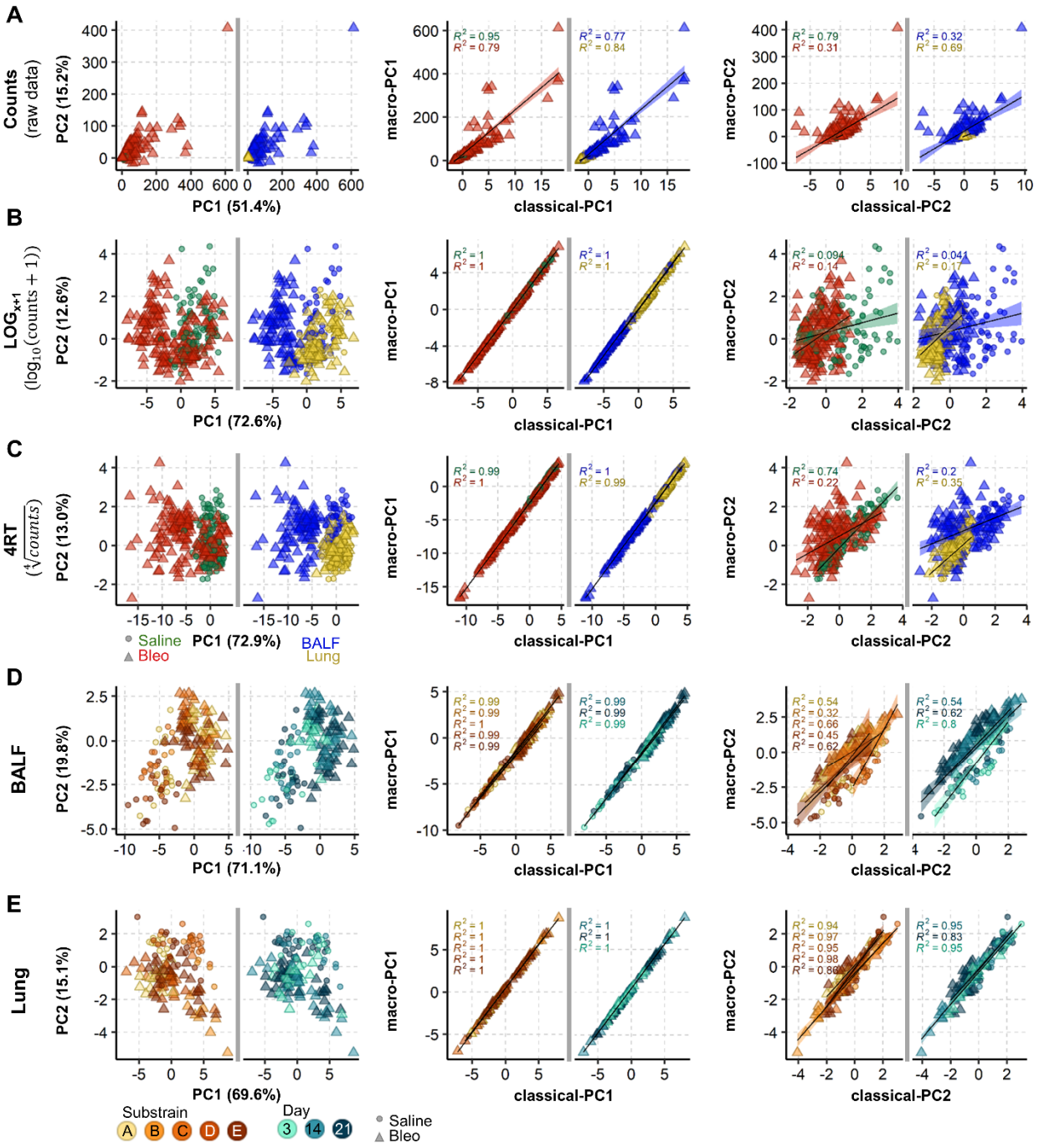

Fig. S2. MacroPCA and PCA deliver similar results. (A-C) MacroPCA scores plot of combined BALF (159 samples) and lung tissue (144 samples), before (untransformed, (A)) and after data transformation by LOG_{x+1} (B) or 4RT (fourth root; (C)). Samples are coloured to highlight effect of bleomycin (Saline or Bleo) and compartment (BALF or Lung). Middle and right panels show the linear fit of the first two principal components derived from the macroPCA and PCA results. (D-E) Separation of entire LOG_{x+1} transformed dataset into the tissue compartments, BALF (D) and lung (E). Middle and right panels show the linear fit of the first two principal components derived from the macroPCA and PCA results. Samples are coloured to highlight different days and substrains. Shapes are in all plots circles for saline and triangles for bleomycin.

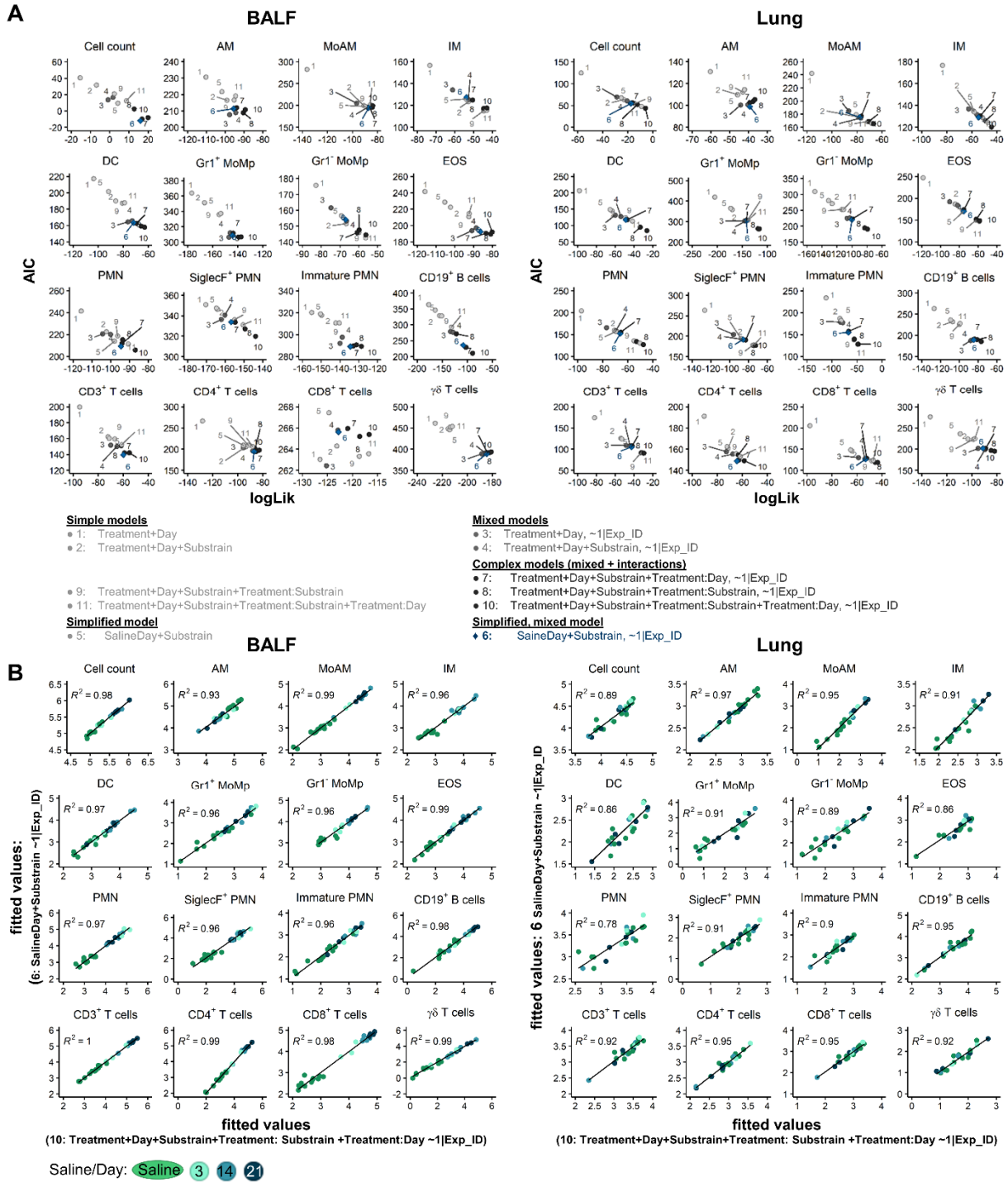

Fig. S3. Simplified mixed models exhibit best performance. Overview of ANOVA model performances for model selection by: (A) Comparison of model performance by AIC and logLik for all 16 cell populations in BALF and lung, better performance is indicated by lower relative estimate of information loss (AIC; Akaike information criterion) and higher goodness of fit (log-likelihood, logLik). (B) Direct comparison of fitted values (on LOG_{x+1} scale) of the simplified mixed model versus the most complex mixed model. The Pearson correlation is shown as black line and R^2 is given.

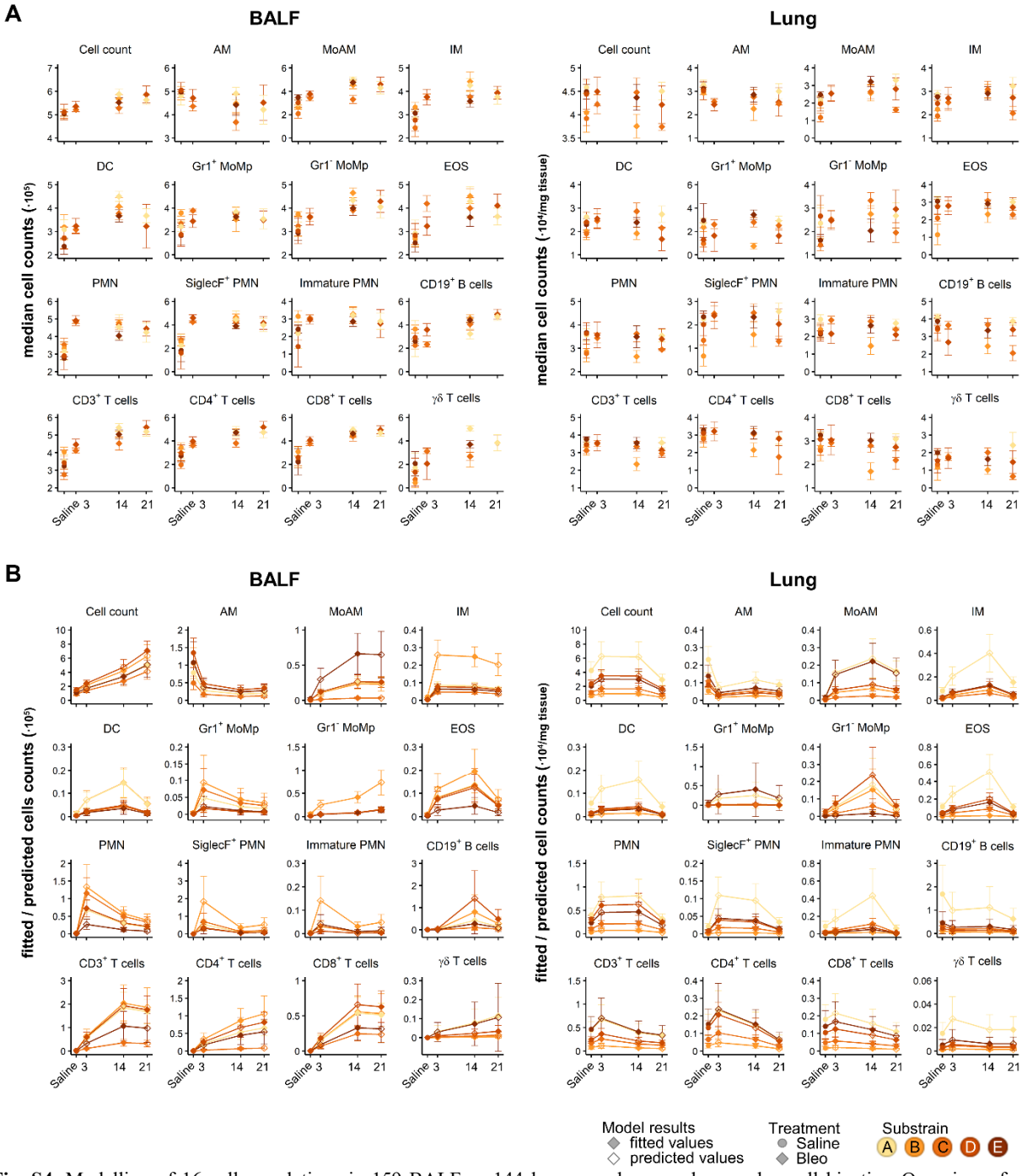

Fig. S4. Modelling of 16 cell populations in 159 BALF or 144 lung samples reveals complex cell kinetics. Overview of ANOVA model performances for model selection by: A) Comparison of model performance by AIC and logLik for all 16 cell populations in BALF and lung, better performance is indicated by lower relative estimate of information loss (AIC; Akaike information criterion) and higher goodness of fit (log-likelihood, logLik). B) Direct comparison of fitted values (on LOG_{x+1} scale) of the simplified mixed model versus the most complex mixed model. The Pearson correlation is shown as black line and R^2 is given.

Supplementary Table S1. Overview of group distribution.

Substrain	A		B		C		D		E	
Compartment	BALF	Lung	BALF	Lung	BALF	Lung	BALF	Lung	BALF	Lung
Condition	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo	Saline | Bleo
									5|8	5|7
Day 3					8|11	8|12			3|4	3|4
			0|4							
Day 14	4|4		0|8	0|9	7|13	4|13	6|10	6|10		
			5|0	7|0						
	5|7								3|6	3|6
Day 21	5|9	5|9				6|3			4|8	4|8
	5|7	5|7								

Supplementary Table S2. Antibodies, fluorophores and sources for flow cytometry.

Panel	Antigen	Label	Company	Catalogue	Clone	Isotype	Identifier	Dilution
Myeloid	CD45	FITC	Thermo Fisher	11-0451-82	30-F11	Rat IgG2b, κ	AB_2753206	1:200
	SiglecF	PE	BD Bioscience	562757	E50-2440	Rat IgG2a, κ	AB_2687994	1:20
	CD11c	ef450	Thermo Fisher	48-0114-82	N418	Armenian hamster IgG	AB_1548654	1:50
	CD11b	ef506	Thermo Fisher	69-0112-82	M1/70	Rat IgG2b, κ	AB_2637406	1:50
	Gr-1 (Ly6G/Ly6C)	PE-Cy7	Biolegend	108402	RB6-8C5	Rat IgG2b, κ	AB_313367	1:800
	CD64a/b	AF647	BD Bioscience	558539	X54-5/7.1	Mouse NOD/Lt IgG1, κ	AB_647120	1:20
	CD24	PerCP Cy5.5	BD Bioscience	562360	M1/69	Rat IgG2b, κ	AB_11151895	1:500
	MHC-II	APC-Cy7	Biolegend	107628	M5/114.15.2	Rat IgG2b, κ	AB_2069377	1:400
Lymphoid	CD45	PerCP Cy5.5	eBioscience	45-0451-82	30-F11	Rat IgG2b, κ	AB_1107002	1:200
	CD3	AF700	Thermo Fisher	56-0033-82	eBio500A2	Syrian hamster / IgG	AB_837094	1:50
	CD19	BB515	BD Bioscience	564531	1D3	Rat IgG2a, κ	AB_2738836	1:50
	CD8	PE	Biolegend	100708	53-6.7	Rat IgG2a, κ	AB_312747	1:100
	CD4	APC	Biolegend	17-0041-82	GK1.5	Rat IgG2b, κ	AB_469320	1:100
	gdTCR	ef450	Thermo Fisher	48-5711-82	eBiogL3	Armenian hamster IgG	AB_2574071	1:50

Supplementary Table S3. Antibodies, fluorophores and sources for immunofluorescent staining.

Antigen	Host	Brand	Catalogue	Identifier	Concentration (µg/ml)
Collagen I	Goat	Southern Biotech	1310-01	AB_2753206	0.8
CD4	Rat	Synaptic Systems	HS-360 017	AB_2800530	10
CD11c	Rabbit	Thermo Fisher	PA5-79537	AB_2746652	3.3
SiglecF	Goat	R&D Systems	AF1706	AB_354943	0.4
Ly6G	Rat	Biolegend	127601	AB_1089179	3.3
CD45	Rabbit	Abcam	AB10558	AB_442810	0.6

Supplementary Table S4. Instrument configurations.

Instrument	Laser lines	Bandpass Filters						
LSRII	488 nm	780/60	695/40	670/14	610/20	576/26	530/30	488/10
	633 nm	780/60	730/45	660/20				
	405 nm	610/20	525/50	440/40				
	355 nm	530/30	440/40					
Cytoflex S	488 nm	690/50	525/40	488/8				
	561 nm	780/60	690/50	610/20	585/42			
	633 nm	780/60	712/25	660/20				
	405 nm	660/20	610/20	525/40	450/45			

Instrument	Parameter	Acquistion seq 1	Acquistion seq 2
Leica TCS-SP8	Pinhole	67.9 µm	67.9 µm
	PinholeAiry	1 AU	1 AU
	EmissionWavelength for PinholeAiry Calculation	580 nm	580 nm
	Excitation Beam Splitter	TD 488/552/638	TD 488/552/638
Hybrid Detectors	HyD 1 (nm)		410 - 460
	HyD 2 (nm)	492 - 522	560 - 571
	HyD 3 (nm)		613 - 630
	HyD 4 (nm)	530 - 548	705 - 740
	HyD 5 (nm)	645 - 675	
Solid state lasers (nm)	405, Intensity (%):	-	0.30
	488, Intensity (%):	0.30	-
	552, Intensity (%):	-	0.40
	638, Intensity (%):	0.30	0.04