ADVANCED THERAPEUTICS
www.advtherap.com

A Spatial Transcriptomics Based Label-Free Method for Assessment of Human Stem Cell Distribution and Effects in a Mouse Model of Lung Fibrosis

Jeongbin Park, Dongjoo Lee, Jae Eun Lee, Daeseung Lee, In Ho Song, Hyun Soo Park, Hongyoon Choi, and Hyung-Jun Im**

Recently, cell therapy has emerged as a promising treatment option for various disorders. Given the intricate mechanisms of action (MOA) and heterogenous distribution in target tissues inherent to cell therapy, it is necessary to develop more sophisticated, unbiased approaches to evaluate the distribution of administered cells and the molecular changes at a microscopic level. This study introduces a label-free approach for assessing the tissue distribution of administered human mesenchymal stem cells (hMSCs) and their MOA, leveraging spatially resolved transcriptomics (ST) analysis. The hMSCs are introduced into a mouse model with lung fibrosis, followed by the manipulation of ST to visualize the spatial distribution of hMSCs within the tissue. This is achieved by capitalizing on interspecies transcript differences between humans and mice. Furthermore, the method allowed for the examination of molecular changes associated with the spatial distribution of hMSCs. Therefore, this method has the potential to serve as an effective tool for various cell-based therapeutic agents.

1. Introduction

Cell therapy involves the introduction of therapeutic cells into patients and has gained significant attention as a promising treatment option for a range of diseases. It has garnered significant attention in recent years as a promising treatment modality for various diseases, including cancers,[1] autoimmune diseases,[2] inflammatory diseases,[3] and neurodegenerative diseases.[4] Cell therapeutic agents are characterized by designability, biocompatibility, and applicability of cell functions even surpassing the blood-brain barrier (BBB),[5] making cell therapy more promising. With the advent of advanced techniques in cellular manipulation (e.g., chimeric antigen receptor-T cell or NK cell therapy (CAR-T/NK), stem cells,

J. Park, D. Lee, J. E. Lee, D. Lee, H. Choi, H.-J. Im
Portrai, Inc.
78-18, Dongsulla-gil, Jongno-gu, Seoul 03136, Republic of Korea
E-mail: chy1000@snu.ac.kr; iiihjjj@snu.ac.kr

I. H. Song, H. S. Park
Department of Nuclear Medicine
Seoul National University College of Medicine
Seoul National University Bundang Hospital
82 Gumi-ro, 173 Beon-gil, Bundang-gu, Seongnam 13620, Republic of Korea

I. H. Song, H. S. Park, H.-J. Im
Department of Molecular Medicine and Biopharmaceutical Sciences
Graduate School of Convergence Science and Technology
Seoul National University
Seoul 08826, Republic of Korea

H. S. Park
Molim, Inc.
Gwanggyo-ro 145, Yeongtong-gu, Suwon 16229, Republic of Korea

H. Choi
Department of Nuclear Medicine
Seoul National University Hospital
Seoul 03080, Republic of Korea

H. Choi
Department of Nuclear Medicine
Seoul National University College of Medicine
Seoul 03080, Republic of Korea

H.-J. Im
Department of Applied Bioengineering
Graduate School of Convergence Science and Technology
Seoul National University
Seoul 08826, Republic of Korea

H.-J. Im
Cancer Research Institute
Seoul National University
Seoul 03080, Republic of Korea

H.-J. Im
Research Institute for Convergence Science
Seoul National University
Seoul 08826, Republic of Korea

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adtp.202300283

DOI: 10.1002/adtp.202300283

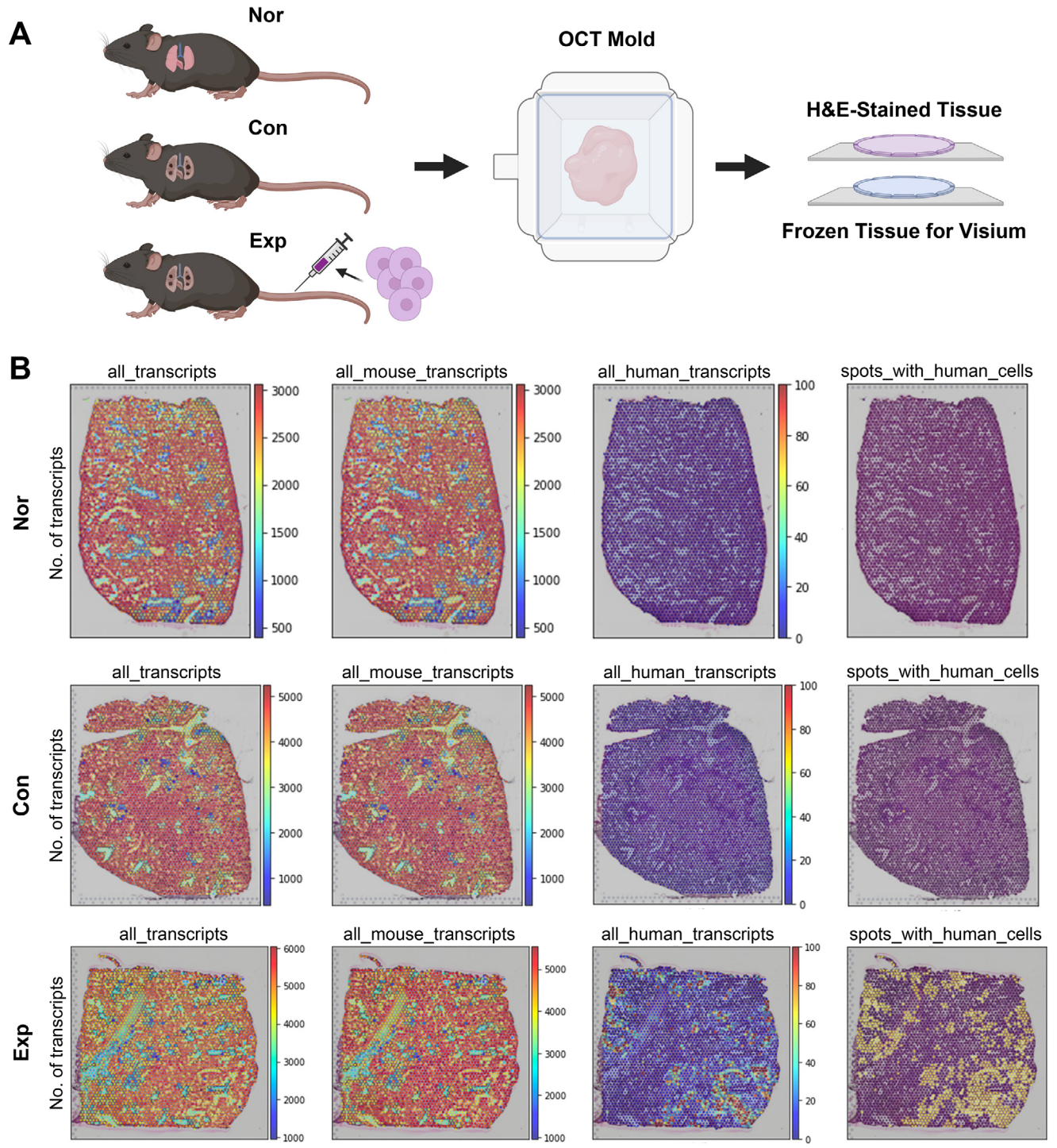

Figure 1. Preparation of lung tissues from normal mouse ("*Nor*"), lung fibrosis model ("*Con*"), and lung fibrosis model treated with hMSCs ("*Exp*"). A) Schematic illustration of ST library preparation. Lung fibrosis and hMSC injection were also represented. B) Number of normalized transcripts according to sample and organism. The results for the "*Nor*", "*Con*", and "*Exp*" samples were represented from top to bottom. The last column of the image, "*spots_with_human_cells*", showed a binary indicator that was yellow only when %human on a spot is greater than the average plus 10 standard deviations of the %human of the "*Con*" sample. Spatial mapping of "*spots_with_human_cells*" well represented the existence of injected human stem cells in the "*Exp*" sample.

Table 1. Specified parameter values for this study.

Function	Parameter	Value
Seurat in R		
FindIntegrationAnchors	*dims*	1:50
IntegrateData	*dims*	1:50
FindNeighbors	*dims*	1:30
FindClusters	*resolution*	0.2
RunUMAP	*dims*	1:30
scanpy in Python		
scanpy.tl.pca	*svd_solver*	arpack
scanpy.pp.neighbors	*n_neighbors*	10
scanpy.pp.neighbors	*n_pcs*	40
scanpy.tl.leiden	*resolution*	0.5
scanpy.pl.rank_genes_groups	*n_genes*	40

tumor-infiltrating lymphocytes (TILs), or microbiome agents), the potential of cell therapy as a game-changing therapeutic approach continues to grow. Despite the promise of cell therapy, it has faced significant challenges due to the living nature of cells, resulting in limitations in reproducibility of drug efficacy and difficulties in preparation, delivery, and administration.[6]

Traditional methods assessing the efficacy of cell therapy often lack the necessary resolution and accuracy to provide meaningful insights into the complex interactions and molecular changes that occur during treatment. For example, polymerase chain reaction (PCR) is performed in organ resolution, limiting the identification of the microscopic properties of cell therapy, particularly in-tissue distribution of administered cells.[7] Fluorescence-labeled cells, which are commonly used to obtain images of administered cells, have a drawback in that the tracers can easily detach from the cells, resulting in limitations for tracking cells.[8] Most of all, it is important to note that these methods do not provide a comprehensive understanding of the detailed molecular changes and interactions that occur between the administered cells and the host cells within the target tissues. These have led to a pressing need for innovative and comprehensive approaches to better understand the true potential of cell therapeutic agents.

Spatially resolved transcriptomics (ST) has emerged as a groundbreaking tool in the field of molecular biology. By combining spatial information with gene expression data, this technique allows researchers to study the dynamic changes in gene expression patterns within tissues and cells at an unprecedented level of detail. ST has the potential to revolutionize the evaluation of cell therapeutic agents by providing a more comprehensive understanding of their effects on the cellular and molecular levels. Because of the capabilities of ST, it was chosen as the Method of the Year in 2020 by Nature Methods.[9] This technology is branching out into areas like sequencing-based ST, image-based ST, and image-based spatial proteomics.[10] Looking ahead, there will be progress in areas such as spatial isoform sequencing by incorporating ST and long-read sequencing,[11] methods for manipulating subcellular information derived from ST,[12] spatially resolved omics studies combining transcriptomes, proteomes, lipidomes, and/or metabolomes,[13] and understanding spatial heterogeneity through approaches like information theory.[14] Additionally, there will be initiatives to broaden the field of view in image-based ST, delve into spatiotemporal sequencing,[15] and explore microbiota[16] and non-mRNAs.[17]

In this paper, we present a novel experimental and analytic procedure that harnesses the power of ST to evaluate cell therapeutic agents. This approach not only offers a more in-depth understanding of the mechanisms underlying the therapeutic effects of these agents but also provides critical insights into the optimization of cell therapy treatments for various diseases. Here, we administered human mesenchymal stem cells (hMSC) in a mouse model of lung fibrosis and presented a method for evaluating both the tissue-level distribution of the administered stem cells and their effect on treated tissue. We expect this analysis method to facilitate the development of cell therapeutics by comprehensively understanding the mode of action and in-depth cell-level distribution in the microenvironment.

2. Results

2.1. Human Transcripts were Analyzed to Evaluate Stem Cell Distribution in a Label-Free Manner

The overall research design and ST library preparation for lung tissues from a normal mouse ("*Nor*"), a control lung fibrosis model ("*Con*"), and a lung fibrosis model treated with hMSCs ("*Exp*") were represented in **Figure 1**A. The parameters and statistical methods were summerized in **Table 1**. Lung fibrosis was induced by intravenous injection of bleomycin for 3 weeks. Lung fibrosis was visually identified on histopathologic images of the lungs (Figure S1, Supporting Information). Normal bone marrow-derived hMSC was intravenously injected in the "*Exp*" sample 6 h before it was sacrificed.

Following the normalization of gene counts, the unreliably large counts of some RNA transcripts were corrected (Figure S2, Supporting Information, the first column). The percentage of the human transcripts of all the transcripts for a single spot, %human, was also explored according to sample and normalization. As a result, %human with normalization was higher in the "*Exp*" sample (0.677 ± 1.911%) compared to the one without normalization (0.417 ± 1.449%) (Figure S2, Supporting Information, the right bottom), while having minimal impact on the "*Nor*" and "*Con*" samples. Also, the precision values for human transcripts improved with normalization (**Tables 2** and **3**). The results demonstrated that the normalization increases the statistical power for detecting human transcripts. Hence, we only considered human transcripts and %human after the normalization process for the subsequent analyses.

We implemented a new alignment method using the bam file from 10x Space Ranger (Figure S3A, Supporting Information). Interestingly, the resultant normalized human counts closely resemble those from Space Ranger, suggesting that the detection of human transcripts is robust regardless of the alignment pipeline. (Figure S3B, Supporting Information). It also suggests that there is a potential to further reduce falsely classified transcripts to the wrong organism by improving the alignment method. The subsequent analyses were performed based on Space Ranger.

We tested various thresholds for %human to address falsely detected human transcripts (Figure S4, Supporting Information).

Table 2. Specificity and precision for detecting human transcripts before normalization.

The "*Nor*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	0	5048
Detected Mouse Transcripts	0	18 406 506
Specificity = 18 406 506/(18 406 506 + 5048) = 0.9997		
The "*Con*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	0	7450
Detected Mouse Transcripts	0	42 812 236
Specificity = 42 812 236/(42 812 236 + 7450) = 0.9998		
The "*Con-2*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	0	7633
Detected Mouse Transcripts	0	27 775 100
Specificity = 27 775 100/(27 775 100 + 7633) = 0.9997		
The "*Exp*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	166 450.5**	11 394.55*
Detected Mouse Transcripts	0 (*assumed*)	37 970 430
Specificity = (0.9997 + 0.9998 + 0.9997)/3 = 0.9997 (*assumed*)		
*37 970 430/(37 970 430 + 11 394.55) = 0.9997		
**166 450.5 = 177 845−11 394.55		
Precision = 166 450.5/177 845 = 0.9359		
The "*Exp-2*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	67 753.28**	6646.718*
Detected Mouse Transcripts	0 (*assumed*)	22 149 080
Specificity = (0.9997 + 0.9998 + 0.9997)/3 = 0.9997 (*assumed*)		
*22 149 080/(22 149 080 + 6646.718) = 0.9997		
**67 753.28 = 74 400−6646.718		
Precision = 67 753.28/74 400 = 0.9107		

Consequently, we defined a threshold based on the average plus 10 times the standard deviation of %human in the "*Con*" sample for two reasons. First, this threshold fell within the range of values for %human where there was no spot exceeding it in the "*Nor*" and "*Con*" samples (Figure S2, Supporting Information, the third column). Second, when applying this threshold, the classified human spots in the "*Exp*" sample displayed a histological structure suitable for comparison in the following cell type mapping results.

The normalized human transcripts were only observable in the "*Exp*" sample compared to the others on a scale from 0 to 100 (Figure 1B, the third column), while low human transcripts were also found in the "*Nor*" and "*Con*" samples in different scales (Figures S2 and S3, Supporting Information). Given that human cells were not administered in the "*Nor*" and "*Con*" samples, the human transcripts detected in the two samples could be considered false positives. However, since the false positive thresholding was effective, the distribution of %human on ST in the "*Exp*" sample can be regarded as the actual distribution of the administered hMSCs (Figure 1B, the fourth column).

2.2. Different Analyses Identified Genes Associated with Administered hMSCs in the Lung Tissue

The spatial clustering analysis was performed only utilizing mouse gene expression, which resulted in 9 clusters (i.e., 0–8) (**Figure 2A**). Among them, cluster 5 showed a significantly higher %human than the others (Figure 2B). In addition, spots from cluster 5 were rarely observed in the "*Nor*" and "*Con*" samples (Figure 2C). It was remarkable that even if human genes were excluded when performing spatial clustering analysis, significantly higher numbers of %human were concentrated on a single cluster. The top 20 spatially enriched mouse genes

Table 3. Specificity and precision for detecting human transcripts after normalization.

The "*Nor*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	0	3176.6821
Detected Mouse Transcripts	0	6 448 516
Specificity = 6 448 516/(6 448 516 + 3176.6821) = 0.9995		
The "*Con*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	0	4954.009
Detected Mouse Transcripts	0	14 301 125
Specificity = 14 301 125/(14 301 125 + 4954.009) = 0.9997		
The "*Con-2*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	0	4819.042
Detected Mouse Transcripts	0	10 613 166
Specificity = 10 613 166/(10 613 166 + 4819.042) = 0.9995		
The "*Exp*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	89 519.62**	5054.463*
Detected Mouse Transcripts	0 (*assumed*)	12 631 102
Specificity = (0.9995 + 0.9997 + 0.9995)/3 = 0.9996 (*assumed*)		
*12 631 102/(12 631 102 + 5054.463) = 0.9996		
**89 519.62 = 94 574.08−5054.463		
Precision = 89 519.62/94 574.08 = 0.9466		
The "*Exp-2*" Sample		
Predicted Class \ True Class	Human Transcripts	Mouse Transcripts
Detected Human Transcripts	41 026.84**	3476.545*
Detected Mouse Transcripts	0 (*assumed*)	8 687 886
Specificity = (0.9995 + 0.9997 + 0.9995)/3 = 0.9996 (*assumed*)		
*8 687 886/(8 687 886 + 3476.545) = 0.9996		
**41 026.84 = 44 503.383−3476.545		
Precision = 41 026.84/44 503.383 = 0.9219		

(SEGs) in cluster 5 are related to immune response, collagen-containing extracellular matrix (ECM), and peptidase activity (Figure 2D).

Differentially expressed genes (DEGs) between the "*Con*" and "*Exp*" samples were obtained. Although we included all genes from humans and mice, the top 6 DEGs were all mouse genes (Figures S5–S7, Supporting Information). To further comprehend genes spatially associated with the hMSC distribution, the spatially associated genes with %human (SAGs) were explored in the "*Exp*" sample by computing Spearman correlation coefficients (**Figure** 3A,B). In this analysis, six genes were human genes among the top 20 SAGs. To assess the molecular process of lung fibrosis in the "*Exp*" sample, we selected positively spatially correlated mouse genes from the top 20 SAGs and performed GO analysis with them. As a result, the top biological pathway was "*transmembrane receptor protein serine/threonine kinase signaling pathway*" (Figure 3C). Also, there were overlapping mouse genes between the SEGs in cluster 5 and the SAGs: *Lcn2* and *Col4a1* (Figures 2 and 3).

2.3. CellDART Revealed Cell types Spatially Associated with hMSC Distribution in the Lung Fibrosis Model

After preparing a single cell RNA sequencing (scRNA-seq) reference of mouse lung (GSE124872), cellular level deconvolution was performed on the "*Exp*" sample using CellDART[18] (**Figure** 4A). Spearman correlation between %human and each CellDART score was calculated in the "*Exp*" sample (Figure 4B). As a result, endothelial cells and epithelial cells were the most and the least associated cell types with %human, respectively. It was well consistent with the fact that the intravenously injected cells were thought to be physically confined in narrow micro-vessels of the alveoli during blood circulation.[19]

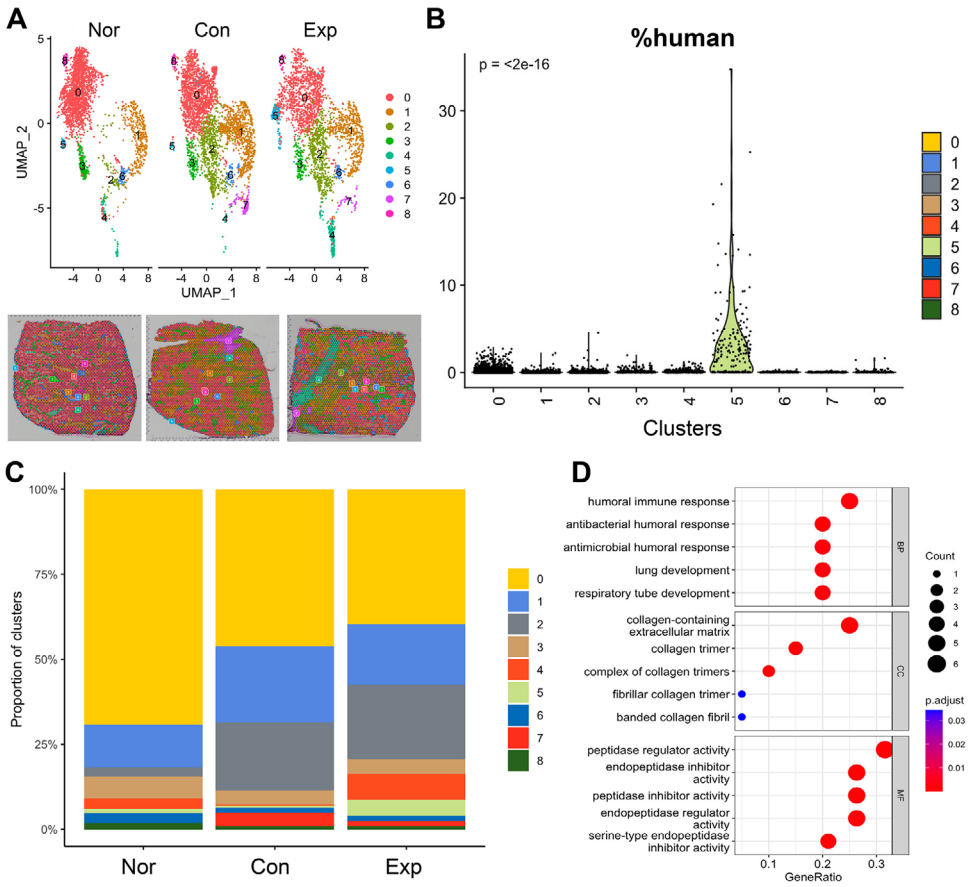

Figure 2. Spatial clustering analysis. A) *DimPlot* (*up*) and *SpatialDimplot* (*down*) of three samples according to clustering labels. B) *VlnPlot* for %human according to clustering labels. Note that the cluster 5 had a notably higher %human compared to the others, even though the clustering was based solely on mouse gene expression. C) The population of clustering labels for each sample. Spots in the cluster 5 rarely appeared in the "*Nor*" and "*Con*" samples. D) GO plot for top 20 SEGs in the cluster 5 (with an adjusted p value of less than 0.05 and ordered by log FC) including *Lcn2* (highest log FC), *Msln*, *Chil3*, *C3*, *Upk3b*, *Spp1*, *Col4a1*, *Serpina3n*, *Wfdc21*, *Col3a1*, *Wfdc17*, *Gm13889*, *Chil1*, *Mgp*, *Sftpd*, *Slpi*, *Ctsc*, *Fmo2*, *Scd1*, and *Napsa* (in order).

We defined the "*is_endothelial*" label for a spot (referred to as an endothelial spot) to indicate whether a spot mainly consists of endothelial cells, using a threshold of the average plus one standard deviation of the endothelial cell score. We then obtained DEGs for the "*Exp*" sample, by comparing the endothelial spots with the "*is_human*" label (referred to as human endothelial spots) to the endothelial spots without that label (referred to as mouse endothelial spots). As a result, only *Apoe*, *Col1a1*, and *Hnrnpab* were significantly enriched (with an adjusted p-value of less than 0.05) in the mouse endothelial spots compared to the human endothelial spots. Among them, *Apoe* and *Col1a1* (activated fibroblast marker) genes are known to be expressed in fibrotic regions surrounding blood vessels.[20] In contrast, *mt-Co2*, *Lcn2*, *Gm42418*, *Chil1*, *Scd1*, *mt-Nd1*, and *mt-Nd2*, along with 43 human genes, were significantly enriched (with an adjusted p-value of less than 0.05) in the human endothelial spots compared to the mouse endothelial spots. Interestingly, *Lcn2*, which is a marker gene for luminal epithelial cells in the mammary gland[21] and goblet cells in the trachea,[22] was found in two previous analyses and this one. This indicates that hMSC may be associated with other cell types that express *Lcn2* as well as with endothelial cells.

2.4. Spatial Cell Colocalization Analysis was Performed Between hMSCs and Mouse Cells

When performing spatial cell colocalization analysis between hMSCs and mouse cells, it was found that endothelial cells exhibited the highest colocalization with %human among six cell types (Jaccard index = 0.304) (**Figure 5**; Figure S10, Supporting Information). This finding aligns with the previous Spearman correlation analysis conducted based on CellDART (Figure 4B). Also, ligand-receptor (LR) interactions between different species were analyzed by using the same analysis and CellTalkDB. To analyze inter-species interactions, CellTalkDB was modified for this study (see Methods). As a result, the *COL1A2* (human) and *Cd93* (mouse) pair exhibited the highest Jaccard index implying the greatest spatial colocalization of the LR pair, where *Cd93* is well known as a marker gene for endothelial cells.[23] In protein alignment studies via BLAST, the homology scores, shown as the percentage of matching sequences, indicate that human *COL1A2* and *CD93* share similarities with mouse *Col1a2* and *Cd93* at 88.89% and 68.24%, respectively. Although additional investigation is needed to confirm inter-species ligand-receptor interactions, this suggests a possibly existent interaction between

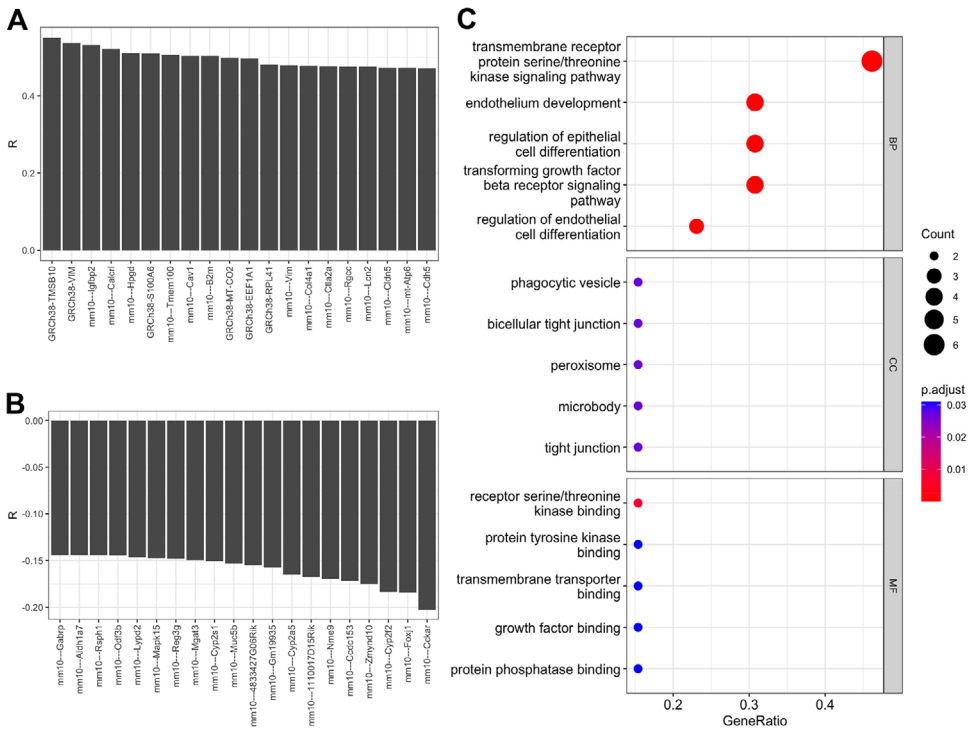

Figure 3. Spatially associated genes with %human (SAGs) in the "*Exp*" sample. The top SAGs and the bottom SAGs were shown in (A,B). C) GO analysis for the mouse genes among top 20 SAGs.

these two genes. This implied a spatial colocalization and potential interaction between mouse endothelial cells and human stem cells.

2.5. Reproducibility Study in a Different Batch of Data

We additionally prepared "*Con-2*" and "*Exp-2*" samples as replicates for the "*Con*" and "*Exp*" samples, respectively. Subsequently, thresholds, parameters, and references were identically applied in the former analyses. As a result, the "*Con-2*" sample showed few human spots like the "*Con*" sample, while the "*Exp-2*" sample exhibited a similar abundance of human transcripts as observed in the "*Exp*" sample (**Figure 6A**). We designated the human transcript as positive and the mouse transcript as negative. Using this classification, we then calculated the specificity with which the human transcript could be detected in three samples that did not contain it: "*Nor*", "*Con*" and "*Con-2*". The specificity values to detect human transcript were found to be 0.9995, 0.9997, 0.9995 for the "*Nor*", "*Con*" and "*Con-2*" samples, respectively. Additionally, if we applied the mean specificity of the three samples to the experimental group and the number of human transcripts misclassified as mouse transcripts is negligible, the precision values were 0.9466 and 0.9219 for the "*Exp*" and "*Exp-2*" samples, respectively. The similarity of these values reflects the reproducibility of human transcripts detection (Table 3).

Moreover, the "*Exp-2*" sample exhibited the highest spatial correlation between hMSCs and mouse endothelial cells, and the lowest with mouse epithelial cells, similar to the "*Exp*" sample. (Figure 6B). Furthermore, both "*Exp*" and "*Exp-2*" samples demonstrated high similarity in the Spearman correlation coefficient values between %human and genes (R-squared = 0.488, Spearman correlation = 0.698) (Figure 6C), as well as between each pair of cell type score and %human (R-squared = 0.836, Spearman correlation = 0.850) (Figure 6D). These findings suggest that analyzing stem cell distribution using ST is reproducible.

3. Discussion

In this paper, we identified three key findings. First, RNA transcripts originating from different genomes can be discriminated and the spatial distribution of them can be acquired. Second, significantly associated genes with the hMSCs can be obtained by spatial clustering analysis, DEG analysis, and SAGs. Lastly, the injected human stem cells displayed a strong association with endothelial cells, while showing an inverse association with epithelial cells.

The market for therapeutic drugs containing nucleic acids is expanding, with numerous drugs currently in development. Clinical trials have been conducted for stem cells,[24] CAR-T therapies,[25] and exosomes containing nucleic acids.[26] Nonetheless, there have been reports of challenges in evaluating the molecular mechanisms of these drugs, particularly interacting with cells in target tissues in preclinical studies. As a result, there is an urgent need for assessment methods for these drugs. Earlier, a method was developed to identify molecular markers spatially associated with an injected drug based on ST.[20] Using this method, markers related to enhanced permeability and retention (EPR) were identified and the method can be applied to

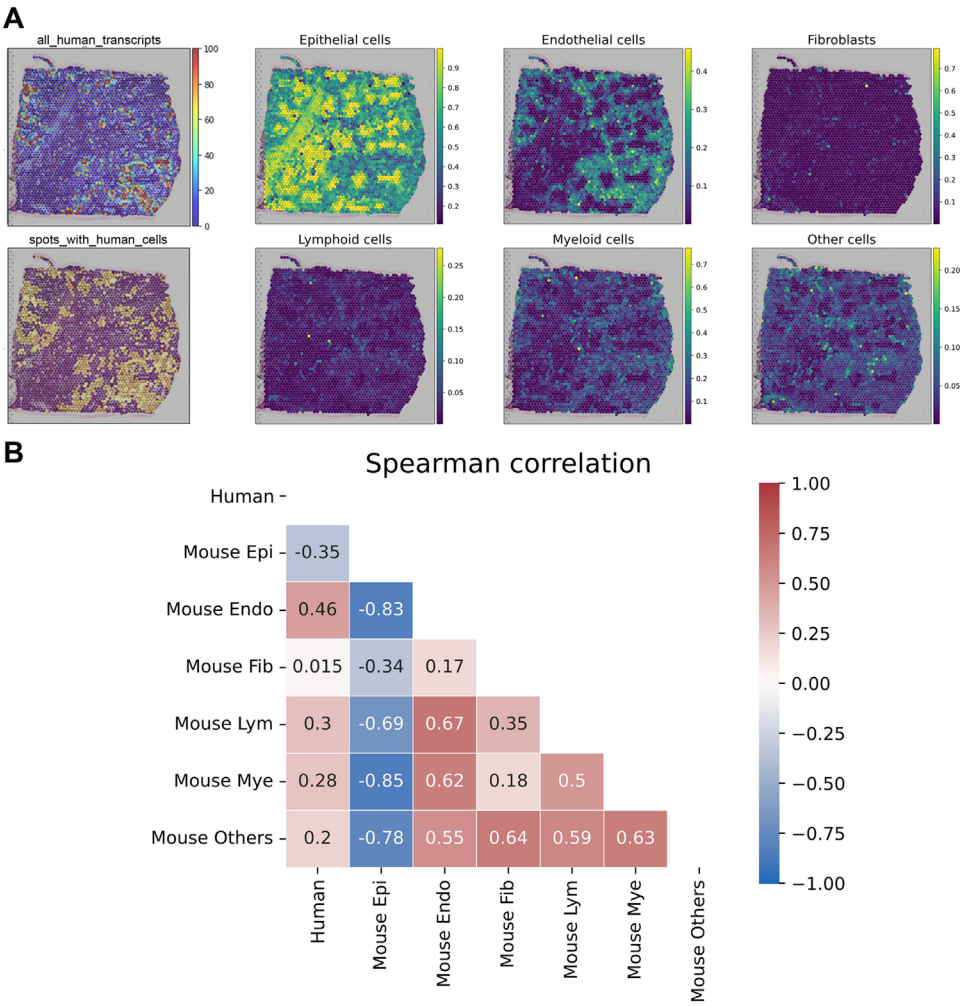

Figure 4. CellDART results. A) Spatial mapping of the proportion of 6 cell types in the "*Exp*" sample, along with human transcripts and %human distribution. B) A correlation coefficient matrix comparing the %human with the CellDART scores for mouse cell types. Specifically, we used the Spearman correlation coefficient for our calculations. Among the cell types examined, endothelial cells showed the highest association with %human, while epithelial cells had the least association.

various therapeutic agents that are labeled with fluorescent dyes. Nonetheless, using dyes to label therapeutics has several drawbacks, including failures to label fluorescent dyes, alteration of properties of drugs, and the possibility that the dyes may detach or disintegrate.[27] The problems were also found in the widely used methods to evaluate cell therapeutic agents using tracer labeling.

This study introduced the spatial mapping of exogenous nucleic acids by using ST in a label-free manner. Previous studies have attempted to map RNA transcripts from different species in applications such as xenografts,[28] host-microbiome mapping,[16] host-virus mapping,[29] and the identification of engineered oligonucleotides.[30] Although these studies demonstrated successful analysis on mixed transcriptomes, they did not employ this separated mapping technique to identify the spatial distribution of cell therapeutics in tissues or estimate their mode of action. In addition, our method can be applied to cell therapeutic agents with the same origin as the host when introducing transfection of genes that do not exist in the host. Our proposed approach has the potential to expand the application of ST for spatial analysis of therapeutics containing exogenous nucleic acids. By using this approach, we can gain a more comprehensive understanding of the mechanisms underlying the therapeutic effects of these agents, which can lead to critical insights for optimizing cell therapy treatments for various diseases.

Apart from simple distribution analysis, spatial transcriptomics analysis of stem cell treatment in our study can reveal the transcriptome-level effects on lung fibrosis tissue. The injected human stem cells showed the up-regulation of hemoglobin genes including *Hba-a1*, *Hba-a2*, *Hbb-bs*, and *Hbb-bt*, which were up-regulated in the "*Exp*" sample compared to the "*Con*" sample (Figures S5 and S7, Supporting Information). In addition, the "*Exp*" sample displayed a decrease in gene expression associated with "*collagen-containing extracellular matrix*" and "*humoral immune response*" (e.g., *Bpifa1*[31]) compared to the "*Con*" sample (Figures S5 and S6, Supporting Information). The molecular changes in the "*Exp*" sample closely mirrored those identified when comparing the "*Nor*" sample with the "*Con*" sample. For

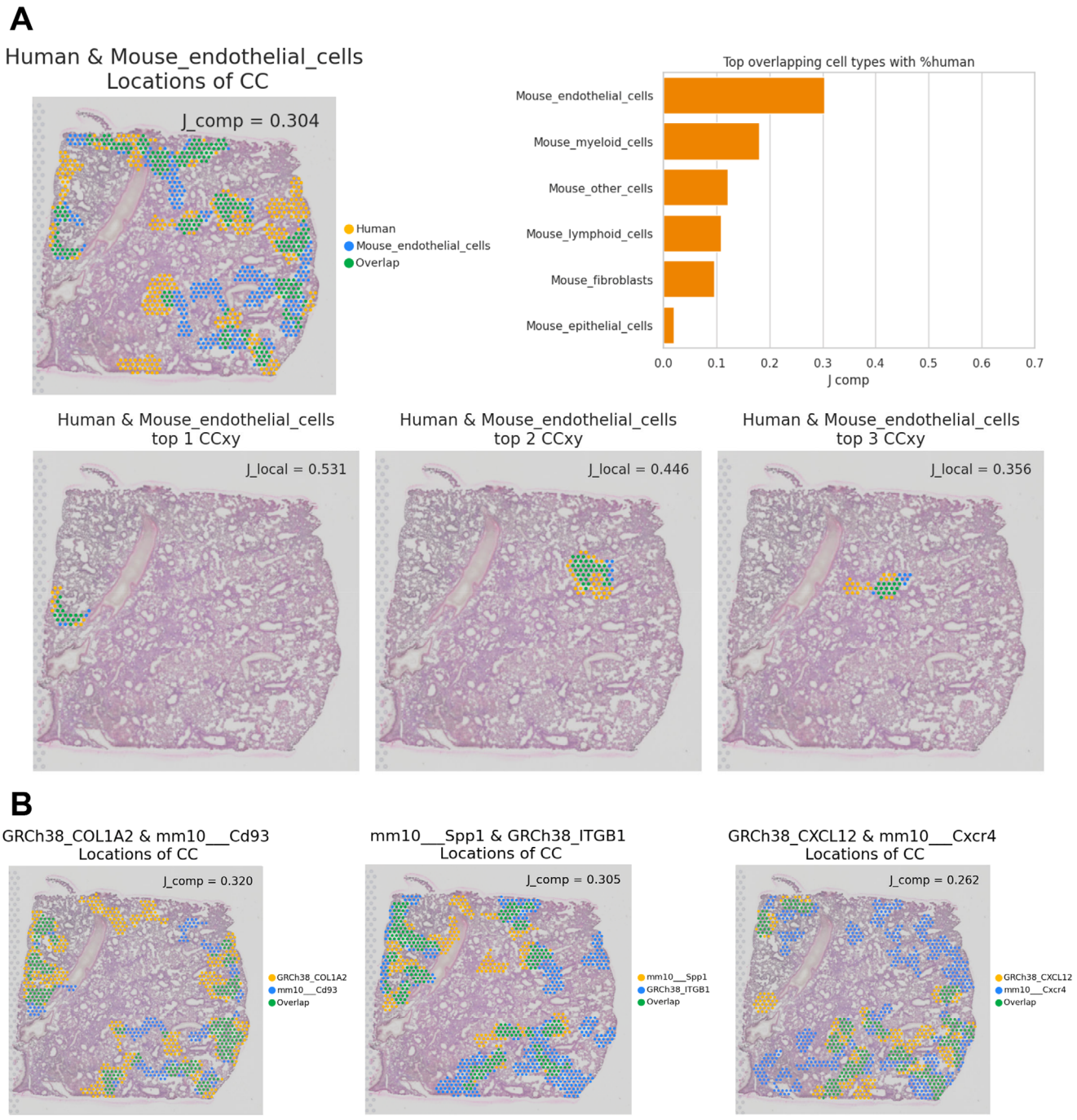

Figure 5. Spatial cell colocalization analysis results. A, The results from spatial cell colocalization analysis comparing the spatial distribution of cell types. Here, we utilized %human and the CellDART scores multiplied by (100 - %human) to indicate the spatial distribution of human and mouse cells. J_comp and J_local on the plots indicate the overall and the local Jaccard index, respectively. Also, CC represents the connected components needed to calculate J_local. B, The topological representation of the LR pairs with the highest Jaccard indices for the LR colocalization. Yellow, blue, and green spots correspond to ligand, receptor, and mixed regions, respectively.

instance, *Hba-a1*, *Hba-a2*, *Hbb-bs*, and *Hbb-bt* were up-regulated, while *Bpifa1* was down-regulated in the "*Nor*" sample when compared with the "*Con*" sample (Figures S8–S9, Supporting Information). It supports the molecular changes were presumed to be the effect of the injected human cells. In addition, the gene expression fold changes in the "*Exp-2*" sample against the "*Con-2*" sample showed high similarity to those in the "*Exp*" sample against the "*Con*" sample (R-squared = 0.469, Spearman correlation = 0.685) (Figure S11, Supporting Information). Still, the comparison of groups at the spot level has limitations because the analysis is only conducted twice after the administration of stem cells. Thus, this study requires using replicate samples and

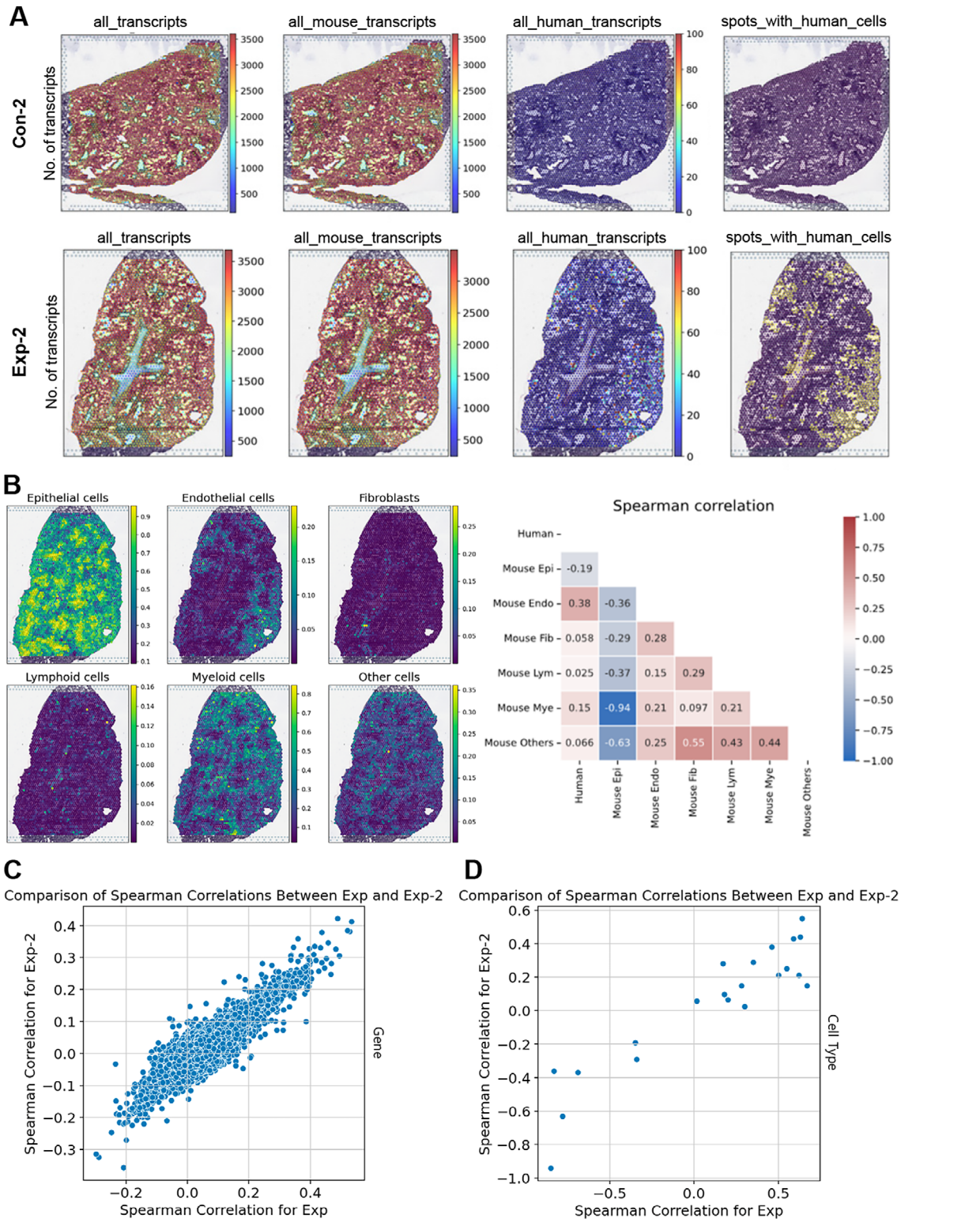

Figure 6. Reproducibility study. A) Number of normalized transcripts according to sample (i.e., "*Con-2*" and "*Exp-2*" samples) and organism. B) Spatial mapping of the proportion of 6 cell types in the "*Exp-2*" sample and a correlation coefficient matrix comparing the %human with the CellDART scores for mouse cell types. C) The correlation coefficients, representing the relationship between %human and gene expression levels, were calculated for each gene across the "*Exp*" and "*Exp-2*" samples. Subsequently, a plot was generated to visually represent this correlation (R-squared values = 0.488, Spearman correlation coefficients = 0.698). D) Similarly, each Spearman correlation coefficient for each pair of cell type score and %human was shown across the "*Exp*" and "*Exp-2*" samples. Afterward, a plot was generated to visually represent this correlation (R-squared values = 0.836, Spearman correlation coefficients = 0.850).

multiple time points to ensure sufficient statistical significance. Also, the estimated MOA effect observed here merely demonstrates the feasibility of this methodological approach and does not necessarily imply causality. To understand the mechanisms of stem cells at the whole transcriptome level, further experiments are needed in addition to in-tissue level distribution analysis. Therefore, it cannot be conclusively regarded as the definitive MOA of hMSC, so additional validation studies are required.

Several challenges should be noted to using spatial analysis of administered cells using ST. The analytic process presented here was dependent on less significant false positive human transcripts in the "*Nor*", "*Con*", and "*Con-2*" samples. The occurrence of false positive human transcripts came from the homology between humans and mice. Also, false negative human transcripts should be addressed. In the top 20 SAGs with %human in the "*Exp*" sample, *VIM* and *Vim* appeared at the same time (Figure 3A). Considering *VIM* is a gene associated with mesenchymal stem cells,[32] it is highly probable that *Vim* counts were falsely classified as mouse transcripts. To resolve the false detection problems presented here, a solution is to adopt long-read sequencing to fully manipulate the genetic sequence differences. Another possible solution is to develop computational algorithms that can manipulate spatial proximity,[33] histological morphology,[34] previous datasets,[18,35] or other domain knowledge to credibly eliminate or modify falsely classified transcripts.

Our study successfully demonstrated the ability of ST to map administered cell therapeutics without the need for fluorescent labeling. As a key result, we showed the spatial distribution of injected human stem cells in lung fibrosis tissue and identified genes associated with the distribution, even though further experiments are needed to fully understand the mechanisms of stem cells in terms of cellular interaction in the target tissue at the whole transcriptome level. Overall, the study highlights the potential of ST for spatial analysis of therapeutics containing exogenous nucleic acids and the need for further research to optimize this approach.

4. Experimental Section

Animal Experiments and Tissue Acquisition: The animal experiments for this study were approved by Seoul National University Bundang Hospital with the IACUC approval code BA-2211-355-002. Normal bone marrow-derived hMSC was prepared from Lonza. We prepared five male C57BL/6 mice, each 9 weeks old. The first class was a normal mouse (labeled "*Nor*"). The second class was two mice with lung fibrosis, serving as a control (labeled "*Con*" or "*Con-2*"). The third class was two mice with lung fibrosis, but they also received an injection of hMSC through their tail veins, making them the experimental group (labeled "*Exp*" or "*Exp-2*"). Lung fibrosis was triggered in both the control ("*Con*" or "*Con-2*") and experimental ("*Exp*" or "*Exp-2*") groups by injecting bleomycin for 3 weeks before they were sacrificed. Bone marrow-derived hMSC was injected intravenously into the "*Exp*" and "*Exp-2*" samples, which were then sacrificed 6 h later.

In determining the optimal time point for the experiment, two critical pieces of physiological background were considered.[36] First, cells, particularly hMSCs, exhibit a rapid pulmonary distribution after administration. Second, intravenously administered cells predominantly enter the right atrium, thereafter circulating through the pulmonary artery and reaching the pulmonary capillaries. Because these cells are about the same size as, or even bigger than, the capillaries, they often remain in the lungs for about two days after being injected. Thus, we chose 6 h, because that's when the pulmonary concentration of hMSCs is expected to be at its highest.

Then, optimal cutting temperature (OCT) blocks (Scigen 4586, USA) were made according to the Visium Spatial Protocols—Tissue Preparation Guide (Document CG000240). The OCT blocks from the experimental group were created 6 h after injecting human stem cells. For each sample of all the classes, two tissue slices were acquired from each OCT mold. One was prepared for H&E staining, and the other was kept as a fresh frozen slice for the Visium ST library.

ST Library Acquisition and Normalization: The tissue sections were fixed, stained, and permeabilized, consulting the Visium Spatial Protocols—Spatial Gene Expression Imaging Guide (Document CG000241), along with tissue optimization (TO) steps. The mRNAs present in the tissues were captured through poly-A tails, and subsequent cDNAs were barcoded and amplified via polymerase chain reaction (PCR) to obtain enough cDNAs for reconstructing libraries. Quantitative PCR (qPCR) and the Agilent Technologies 4200 TapeStation were used to measure and assess the quality of the libraries, respectively. Finally, the libraries were sequenced using Illumina HiSeq 4000, following the instructions provided in the user guide. We also used *spaceranger mkref* in the Space Ranger (version 2.0.1) pipeline to combine the human reference, *GRCh38*, with the mouse reference, *mm10*. After that, we ran *spaceranger count* on each sample to generate the processed ST cDNA library against a combined human and mouse reference genome to detect human transcripts from the administered cells in mouse tissues.

The raw count of a gene in a spot does not necessarily reflect the biological activity of the gene in that spot, as it can be influenced by the sequencing depth. To address this, gene counts are adjusted using various methods, such as the trimmed mean of M-values (TMM).[37] Also, since genes with large counts can overestimate their activities, gene counts are generally log-transformed. These processes, referred to as normalization, can be performed using *scanpy.pp.normalize_total* and *scanpy.pp.log1p* functions in Python.

The genomic data stored in BAM format was processed per individual barcode for an alternative alignment. For each barcode, reads from the main BAM file *possorted_genome_bam.bam*, generated by *spaceranger count*, were filtered based on the specified barcode using *samtools view* from SAMtools (version 1.10-3). Filtered reads were then sorted by their names, and unmapped reads were subsequently removed. The remaining reads were sorted by their genomic coordinates and indexed. Read counts per gene were then computed using *htseq-count* from HTSeq (version 2.0.4) with gene annotations provided by the GTF file in the combined reference. The processing for each barcode was carried out concurrently using GNU Parallel to leverage multiple CPU cores for improved computational efficiency.

Acquisition of DEGs Related to hMSC Distribution: Data integration was performed with *FindIntegrationAnchors* and *IntegrateData* in the Seurat (ver. 4.3.0) package. Then, spatial clustering analysis was performed after eliminating human genes from the integrated Seurat object. The elimination process was thought to be meaningful not to make biases toward the "*Exp*" sample which only contained true human transcripts. However, the elimination process was not performed in the other analyses. The top 20 spatially enriched mouse genes (SEGs) (with an adjusted p-value of less than 0.05 and ordered by log FC) were identified from a specific cluster containing most human RNA transcripts using the Wilcoxon Rank Sum test. This was achieved by the *FindAllMarkers* function in Seurat. We then used the *enrichGO* function from the clusterProfiler package (ver. 4.6.2) to analyze these genes in a gene ontology (GO) plot. These plots were categorized into five terms based on three groups: biological process (BP), cellular component (CC), and molecular function (MF).

To analyze differences between samples, we identified differentially expressed genes (DEGs) using the Wilcoxon Rank Sum test with the *FindAllMakers* function in Seurat. We then examined the top 20 DEGs (with an adjusted p-value of less than 0.05 and ordered by log FC) in GO analysis. To find spatially associated genes with %human (SAGs), we computed the Spearman correlation coefficients between %human and gene expression, but only within each experimental sample. Subsequently, we investigated

the mouse genes among the top 20 SAGs (with an adjusted p-value of less than 0.05 and ordered by Spearman correlation) in GO analysis.

CellDART: To identify spatially associated cell types with the distribution of %human, we performed cell type inference by domain adaptation of single-cell and spatial transcriptomic data (CellDART).[18] Here, the single-cell RNA sequencing (scRNA-seq) reference was created by utilizing a publicly available mouse lung scRNA-seq reference (GSE124872). During this process, 33 cell types were collapsed into 7 coarse cell types, and cell types that were ambiguous to classify were categorized as unknown, which was excluded from subsequent processes. Spearman correlation coefficients between CellDART scores of mouse cell types and %human were calculated, only focusing on each experimental sample.

STopover: One of the main therapeutic mechanisms of administered stem cells arises from various paracrine effects stemming from cell–cell interactions.[38] When interactions are distinctly different between cell types, the mechanism of cell therapeutic agents can be more easily discerned. Thus, analyzing cell–cell interactions between different biological systems, such as mouse and human cells, offers valuable insights into the mechanism of injected cell therapeutic agents. Several algorithms for analyzing cell–cell interactions have been developed including CCCExplorer, CellChat, ICELLNET, and CellPhoneDB.[39] Nevertheless, many of these algorithms overlook the importance of spatial proximity. However, recent studies have attempted to address this issue by employing machine learning algorithms on ST data.[40] In line with these efforts, our team has also developed STopover, an algorithm that leverages topological analysis techniques to estimate the co-enrichment of two variables (e.g., cell type score, gene expression, or %human) in ST data.[41]

This algorithm computes the Jaccard index, which evaluates $|A \cap B|$ divided by $|A \cup B|$, given sets A and B. The index can be calculated globally (J_comp) or locally (J_local), with the latter calculated separately according to the connected components (CCs). The parameters of STopover were set by default. Also, the spatial association of cell types was explored by using STopover. Here, we used %human and the CellDART scores multiplied by (100 - %human) to represent the spatial distribution of human and mouse cells, respectively.

After that, ligand-receptor interactions between different species were explored. Existing ligand-receptor databases, including CellPhoneDB[42] and CellTalkDB,[43] are typically focused on individual species, primarily human or mouse. The study of ligand-receptor interactions in different species has been an unexplored area. To do this, we identified the shared ligand-receptor (LR) pairs between humans and mice in CellTalkDB. Then, we used the human and mouse gene homology database derived from the Mouse Genome Informatics (MGI) database to convert the shared LR pairs. These pairs were replaced by two different types of pairs: one involving a human ligand with a mouse receptor, and the other involving a mouse ligand with a human receptor. Then, the colocalization of these LR pairs was evaluated using STopover based on their normalized gene expressions in *Scanpy*, Python. Finally, the mouse-human homology score for each gene of the LR pairs was calculated by referring to pairwise protein alignments using BLAST at NCBI HomoloGene.

Statistical Analysis: R (ver 4.0.5) and Python (ver 3.7.12) were used as programming languages. In addition, *Seurat* (ver 4.3.0), *scanpy* (ver 1.9.1), Space Ranger (ver 2.0.1), SAMtools (ver 1.10-3), and HTSeq (ver 2.0.4) were used. *GRCh38* (*Homo sapiens*) and *mm10* (*Mus musculus*) were regarded as references for Space Ranger and custom alignment. When presenting specific statistics, they were displayed as mean ± SD, rank-based correlation coefficient, R-squared, or Jaccard index, and were distinguished as p-value and adjusted p-value depending on whether they underwent multiple hypothesis testing. DEGs were explored by sorting against fold change (FC) values for all genes with adjusted p values less than 0.05. When drawing GO plots, 20 genes were selected and analyzed. The parameters and statistical methods were set to their default values unless otherwise specified in Table 1,3.

Ethics Approval Statement: Animal experiments for this study were approved by the Seoul National University Bundang Hospital of the Republic of Korea with the approval code of BA-2211-355-002. There were no other ethical issues for this study.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This research received funding from the National Research Foundation of Korea (NRF) under the grant numbers NRF- 2020M3A9B6038086, NRF- 2023R1A2C2006636, NRF- 2022M3A9D3016848, NRF-2020R1C1C1009000, and NRF-2021M2E8A1039564. Also, it was supported by a grant from the Korea Evaluation Institute of Industrial Technology (KEIT) funded by the Korean government (MOTIE) (No. 20018522) and by the Korea Drug Development Fund funded by the Ministry of Science and ICT, Ministry of Trade, Industry and Energy and the Ministry of Health and Welfare (HN22C0632) of Korea.

Conflict of Interest

D.L., H.C., and H.-J.I. are the co-founders of Portrai, Inc, while H.S.P. is a founder of Molim, Inc. Schematic illustrations were generated by BioRender under a complete license. Otherwise, there is no competing financial interest.

Author Contributions

H.C. and H.-J.I. designed this project. Preparation of mouse lung fibrosis model, acquisition of lung tissues and tissue optimization were done by J.E.L., I.H.S., and H.S.P. The library preparation for spatial transcriptomics were performed by J.E.L. J.P. and D.L. ran downstream pipelines and bioinformatics analyses. D.S.L., H.C., and H.J.I. performed data analysis and interpretation. The manuscript was written by J.P., H.C., and H.J.I. and all authors have contributed to the completion of the manuscript.

Data Availability Statement

The raw and processed Visium data from five mice tissues will be available in Gene Expression Omnibus (GEO) under the accession number GSE253378, subject to an embargo period starting from the publication date.

Keywords

cell therapy, interspecies transcript difference, label-free approach, mechanisms of action, spatial distribution, spatially resolved transcriptomics

Received: September 13, 2023
Revised: January 16, 2024
Published online: February 6, 2024

[1] A. R. Saez-Ibanez, S. Upadhaya, T. Partridge, M. Shah, D. Correa, J. Campbell, *Nat. Rev. Drug Discov.* **2022**, *21*, 631.
[2] F. Ghobadinezhad, N. Ebrahimi, F. Mozaffari, N. Moradi, S. Beiranvand, M. Pournazari, F. Rezaei-Tazangi, R. Khorram, M. Afshinpour, R. A. Robino, A. R. Aref, L. M. R. Ferreira, *Front Immunol.* **2022**, *13*, 1075813.
[3] N. Hossein-Khannazer, S. Torabi, R. Hosseinzadeh, S. Shahrokh, H. Asadzadeh Aghdaei, A. Memarnejadian, N. Kadri, M. Vosough, *Hum Cell* **2021**, *34*, 1289.

[4] A. U. Pradhan, O. Uwishema, H. Onyeaka, I. Adanur, B. Dost, *Brain Behav.* **2022**, *12*, 2740.
[5] Y. Xia, L. Rao, H. Yao, Z. Wang, P. Ning, X. Chen, *Adv. Mater.* **2020**, *32*, 2002054.
[6] W. Z. Zhuang, Y. H. Lin, L. J. Su, M. S. Wu, H. Y. Jeng, H. C. Chang, Y. H. Huang, T. Y. Ling, *J. Biomed. Sci.* **2021**, *28*, 28.
[7] A. Brooks, K. Futrega, X. Liang, X. Hu, X. Liu, D. H. G. Crawford, M. R. Doran, M. S. Roberts, H. Wang, *Stem Cells Transl. Med.* **2018**, *7*, 78.
[8] W. G. Kreyling, A. M. Abdelmonem, Z. Ali, F. Alves, M. Geiser, N. Haberl, R. Hartmann, S. Hirn, D. J. de Aberasturi, K. Kantner, G. Khadem-Saba, J. M. Montenegro, J. Rejman, T. Rojo, I. R. de Larramendi, R. Ufartes, A. Wenk, W. J. Parak, *Nat. Nanotechnol.* **2015**, *10*, 619.
[9] V. Marx, *Nat. Methods* **2021**, *18*, 9.
[10] a) J. R. Moffitt, E. Lundberg, H. Heyn, *Nat. Rev. Genet.* **2022**, *23*, 741; b) L. Moses, L. Pachter, *Nat. Methods* **2022**, *19*, 534.
[11] E. Boileau, X. Li, I. S. Naarmann-de Vries, C. Becker, R. Casper, J. Altmuller, F. Leuschner, C. Dieterich, *Front Genet.* **2022**, *13*, 912572.
[12] a) Z. Fang, A. J. Ford, T. Hu, N. Zhang, A. Mantalaris, A. F. Coskun, *Cell Rep. Methods* **2023**, *3*, 100476; b) A. Kumar, A. W. Schrader, A. E. Boroojeny, M. Asadian, J. Lee, Y. J. Song, S. D. Zhao, H. S. Han, S. Sinha, *Res. Sq.* **2023**.
[13] a) C. Sun, A. Wang, Y. Zhou, P. Chen, X. Wang, J. Huang, J. Gao, X. Wang, L. Shu, J. Lu, W. Dai, Z. Bu, J. Ji, J. He, *Nat. Commun.* **2023**, *14*, 2692; b) Z. Miao, B. D. Humphreys, A. P. McMahon, J. Kim, *Nat. Rev. Nephrol.* **2021**, *17*, 710.
[14] a) A. Kashyap, M. A. Rapsomaniki, V. Barros, A. Fomitcheva-Khartchenko, A. L. Martinelli, A. F. Rodriguez, M. Gabrani, M. Rosen-Zvi, G. Kaigala, *Trends Biotechnol.* **2022**, *40*, 647; b) T. Vo, B. Balderson, K. Jones, G. Ni, J. Crawford, A. Millar, E. Tolson, M. Singleton, M. Kojic, T. Robertson, S. Walters, O. Mulay, D. D. Bhuva, M. J. Davis, B. J. Wainwright, Q. Nguyen, L. A. Genovesi, *Genome Med.* **2023**, *15*, 29.
[15] P. Wahle, G. Brancati, C. Harmel, Z. He, G. Gut, J. S. Del Castillo, A. Xavier da Silveira Dos Santos, Q. Yu, P. Noser, J. S. Fleck, B. Gjeta, D. Pavlinic, S. Picelli, M. Hess, G. W. Schmidt, T. T. A. Lummen, Y. Hou, P. Galliker, D. Goldblum, M. Balogh, C. S. Cowan, H. P. N. Scholl, B. Roska, M. Renner, L. Pelkmans, B. Treutlein, J. G. Camp, *Nat. Biotechnol. 41*, **2023**, 1765.
[16] J. L. Galeano Nino, H. Wu, K. D. LaCourse, A. G. Kempchinsky, A. Baryiames, B. Barber, N. Futran, J. Houlton, C. Sather, E. Sicinska, A. Taylor, S. S. Minot, C. D. Johnston, S. Bullman, *Nature* **2022**, *611*, 810.
[17] D. W. McKellar, M. Mantri, M. M. Hinchman, J. S. L. Parker, P. Sethupathy, B. D. Cosgrove, I. De Vlaminck, *Nat. Biotechnol.* **2023**, *41*, 513.
[18] S. Bae, K. J. Na, J. Koh, D. S. Lee, H. Choi, Y. T. Kim, *Nucleic Acids Res.* **2022**, *50*, 57.
[19] K. O. Jung, T. J. Kim, J. H. Yu, S. Rhee, W. Zhao, B. Ha, K. Red-Horse, S. S. Gambhir, G. Pratx, *Nat. Biomed. Eng.* **2020**, *4*, 835.
[20] J. Park, J. Choi, J. E. Lee, H. Choi, H. J. Im, *Small Methods* **2022**, *6*, 2201091.
[21] O. Franzen, L. M. Gan, J. L. M. Bjorkegren, *Database* **2019**, *2019*, baz046.
[22] C. Tabula Sapiens, R. C. Jones, J. Karkanias, M. A. Krasnow, A. O. Pisco, S. R. Quake, J. Salzman, N. Yosef, B. Bulthaup, P. Brown, W. Harper, M. Hemenez, R. Ponnusamy, A. Salehi, B. A. Sanagavarapu, E. Spallino, K. A. Aaron, W. Concepcion, J. M. Gardner, B. Kelly, N. Neidlinger, Z. Wang, S. Crasta, S. Kolluru, M. Morri, A. O. Pisco, S. Y. Tan, K. J. Travaglini, C. Xu, M. Alcantara-Hernandez, et al., *Science* **2022**, *376*, eabl4896.
[23] a) A. Ianevski, A. K. Giri, T. Aittokallio, *Nat. Commun.* **2022**, *13*, 1246; b) A. De Falco, F. Caruso, X. D. Su, A. Iavarone, M. Ceccarelli, *Nat. Commun.* **2023**, *14*, 1074.
[24] Y. Ji, C. Hu, Z. Chen, Y. Li, J. Dai, J. Zhang, Q. Shu, *Stem Cell Res. Ther.* **2022**, *13*, 307.
[25] N. A. Ivica, C. M. Young, *Healthcare* **2021**, *9*, 1062.
[26] Y. Zhang, Q. Liu, X. Zhang, H. Huang, S. Tang, Y. Chai, Z. Xu, M. Li, X. Chen, J. Liu, C. Yang, *J. Nanobiotechnol.* **2022**, *20*, 279.
[27] a) A. K. Srivastava, J. W. Bulte, *Stem Cell Rev. Rep.* **2014**, *10*, 127; b) P. K. Horan, M. J. Melnicoff, B. D. Jensen, S. E. Slezak, *Methods Cell Biol.* **1990**, *33*, 469; c) H. Choi, D. S. Lee, *Stem Cell Res. Ther.* **2016**, *7*, 55.
[28] Z. Ni, A. Prasad, S. Chen, R. B. Halberg, L. M. Arkin, B. A. Drolet, M. A. Newton, C. Kendziorski, *Nat. Commun.* **2022**, *13*, 2971.
[29] H. Sounart, E. Lázár, Y. Masarapu, J. Wu, T. Várkonyi, T. Glasz, A. Kiss, E. Borgström, A. Hill, SA. Jurek, A. Niesnerova, H. Druid, O. Bergmann, S. Giacomello, *Genome Biology* **2023**.
[30] N. Ben-Chetrit, X. Niu, A. D. Swett, J. Sotelo, M. S. Jiao, C. M. Stewart, C. Potenski, P. Mielinis, P. Roelli, M. Stoeckius, D. A. Landau, *Nat. Biotechnol.* **2023**, *41*, 788.
[31] Y. A. Tsou, M. C. Tung, K. A. Alexander, W. D. Chang, M. H. Tsai, H. L. Chen, C. M. Chen, *Biomed Res. Int.* **2018**, *2018*, 1.
[32] a) L. Zhang, Y. Wei, Y. Chi, D. Liu, S. Yang, Z. Han, Z. Li, *Cell Biosci.* **2021**, *11*, 6; b) T. T. Cooper, S. E. Sherman, G. I. Bell, J. Ma, M. Kuljanin, S. E. Jose, G. A. Lajoie, D. A. Hess, *Stem Cells* **2020**, *38*, 666.
[33] Y. Long, K. S. Ang, M. Li, K. L. K. Chong, R. Sethi, C. Zhong, H. Xu, Z. Ong, K. Sachaphibulkij, A. Chen, L. Zeng, H. Fu, M. Wu, L. H. K. Lim, L. Liu, J. Chen, *Nat. Commun.* **2023**, *14*, 1155.
[34] a) C. Comiter, E. D. Vaishnav, M. Ciapmricotti, B. Li, Y. Yang, S. J. Rodig, M. Turner, K. L. Pfaff, J. Jané-Valbuena, M. Slyper, *bioRxiv* **2023**, 2023; b) M. Parreno-Centeno, G. Malagoli Tagliazucchi, E. Withnell, S. Pan, M. Secrier, bioRxiv **2022**, 2022.
[35] a) H. Li, J. Zhou, Z. Li, S. Chen, X. Liao, B. Zhang, R. Zhang, Y. Wang, S. Sun, X. Gao, *Nat. Commun.* **2023**, *14*, 1548; b) B. F. Miller, F. Huang, L. Atta, A. Sahoo, J. Fan, *Nat. Commun.* **2022**, *13*, 2339.
[36] a) A. Bansal, M. K. Pandey, Y. E. Demirhan, J. J. Nesbitt, R. J. Crespo-Diaz, A. Terzic, A. Behfar, T. R. DeGrado, *EJNMMI Res.* **2015**, *5*, 19; b) D. L. Kraitchman, M. Tatsumi, W. D. Gilson, T. Ishimori, D. Kedziorek, P. Walczak, W. P. Segars, H. H. Chen, D. Fritzges, I. Izbudak, R. G. Young, M. Marcelino, M. F. Pittenger, M. Solaiyappan, R. C. Boston, B. M. Tsui, R. L. Wahl, J. W. Bulte, *Circulation* **2005**, *112*, 1451.
[37] M. D. Robinson, A. Oshlack, *Genome Biol.* **2010**, *11*, R25.
[38] a) X. Liang, Y. Ding, Y. Zhang, H. F. Tse, Q. Lian, *Cell Transplant.* **2014**, *23*, 1045; b) M. Gnecchi, Z. Zhang, A. Ni, V. J. Dzau, *Circ. Res.* **2008**, *103*, 1204.
[39] E. Armingol, A. Officer, O. Harismendy, N. E. Lewis, *Nat. Rev. Genet.* **2021**, *22*, 71.
[40] D. S. Fischer, A. C. Schaar, F. J. Theis, *Nat. Biotechnol.* **2023**, *41*, 332.
[41] S. Bae, H. Lee, K. J. Na, D. S. Lee, H. Choi, Y. T. Kim, *bioRxiv* **2022**, *2022*, 11.
[42] M. Efremova, M. Vento-Tormo, S. A. Teichmann, R. Vento-Tormo, *Nat. Protoc.* **2020**, *15*, 1484.
[43] X. Shao, J. Liao, C. Li, X. Lu, J. Cheng, X. Fan, *Brief Bioinform.* **2021**, *22*, bbaa269.