DISSERTATION

A QUPATH WORKFLOW UTILIZING MACHINE LEARNING TO ANALYZE HOMING PROTEIN SPECIFICITY AND PENETRATION INTO LUNG GRANULOMAS OF *MYCOBACTERIUM TUBERCULOSIS* INFECTED MICE

Submitted by

John Patterson

Department of Microbiology Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2024

Master Committee:

Advisor: Mercedes Gonzalez-Juarrero

Mike Lyons
Soham Ghosh

ABSTRACT

A QUPATH WORKFLOW UTILIZING MACHINE LEARNING TO ANALYZE HOMING PROTEIN SPECIFICITY AND PENETRATION INTO LUNG GRANULOMAS OF *MYCOBACTERIUM TUBERCULOSIS* INFECTED MICE

Targeted delivery of drugs to the lungs can improve TB chemotherapy and thus our goal is to develop TB-drug loaded nanoparticles tagged to pulmonary homing peptides. In a previous study, homing peptides to the lungs of TB diseased animals were identified using preclinical TB models (Balb/c and C3HeB/FeJ mice). The selection of homing peptides was carried out using a phage library containing peptides with known homing affinity in other diseases (e.g. cancer). Having identified and selected the homing peptide PL1 (PPRRGLIKLKTS) to granulomas present in the lungs of murine TB models, the PL1 peptide and a negative control (scrambled LinnTT peptide) were tagged to Fluorescein Amidites (FAM). To facilitate tracking *in vivo* of the nanoparticles to be loaded with TB drugs, silver nanoparticles (SNP) were conjugated to Cy3 fluorochrome, a fluorescent marker used in *in vivo* tracking studies, followed by functionalization with the PL1 homing peptide (PL1-SNP) or biotin as negative control (Ctrl-SNP). Tracking and homing of the PL1 peptide to granulomas was possible after *in vivo* administration via intraperitoneal (IP) or intravenous (IV) route of either the FAM tagged synthetic peptides or Cy3-SNPs to *Mycobacterium tuberculosis* (Mtb) infected C3HeB/FeJ mice. Visualization of the fluorescence-tagged carriers within the lungs was performed using microscopic slides affixed with lung sections from each mouse followed by whole slide imaging. The semi-quantitative analysis of the fluorescence whole slide images performed using the QuPath workflow confirmed that PL1-FAM, or PL1-SNP homed to the granulomas. Thereafter, a QuPath workflow

was developed that uses machine learning approaches (MLP) for unbiased identification of tissue types. Other tools were used for characterization and quantification of FAM (synthetic peptides) and Cy3 (SNP) positive cells within granulomatous lesions of the C3HeB/FeJ TB mouse model. Moreover, it was important to quantify the penetration capacity of the FAM tagged peptide as well as the peptide coated SNP into granulomas. QuPath also includes a built in MLP pixel classifier for unbiased segmentation of the whole slide. In addition, a modified QuPath script was developed to segment the granulomas into concentric regions (outer, inner and center) followed by detection and quantification of positive cells for either fluorochrome within each region. Specific colocalization of PL1 with its known receptor (FN-EDB), either as a synthetic peptide or coupled to the SNP, was also studied using lung sections from mice treated with PL1-FAM or PL1-SNP and counter stained with Alexa 647 conjugated anti-FN-EDB monoclonal antibodies. The modified QuPath script was trained to quantify fluorescence from Alexa 647 in cells within granulomas and the Pearson coefficient and QuPath script was used to assess PL1 and FN-EDB colocalization within each region of the granuloma. The results demonstrated that when compared to their respective control samples, the IP route of administration provides equal or better homing of PL1 peptide to the granulomas than the IV route. Both the PL1-FAM and PL1-SNP home to the granulomas and specifically colocalize with its receptor FN-EDB. The FAM tagged peptide and SNP penetrate to the inner and center regions of the granuloma whereas the control SNP were unable to penetrate the barrier in the outer region of the granulomas. The QuPath workflow developed here can be used for tracking and quantification of other homing peptides and nanoparticles for development of new TB therapeutics.

ACKNOWLEDGMENTS

I would first like to thank my advisor, Mercedes Gonzalez-Juarrero, who has provided the encouragement and support necessary to complete this thesis. Her patience and understanding have made it a pleasure to have her as my advisor. Her assistance in the writing process and interpretation of data has played a critical role in completing this research project.

I would also like to show my appreciation for my committee members, Soham Ghosh, and Michael Lyons. Thier expertise has been crucial in the development of the research project and in keeping me focused. Their invaluable suggestions helped reassure me that I was up to the task.

In addition, this project would not have been possible without the guidance and the research done by Greg Robertson and Tambet Teesalu. I would also like to acknowledge the hard work done Maaarja Haugas and Allison Bauman, who provided me with the samples needed to for all the analysis. I would also like to thank all of those involved in the development of the homing nanoparticles used in this research.

I would like to express my gratitude to Peter Bankhead, PhD, University of Edinburgh, for developing the opensource histopathology software, Qupath. Furthermore, this analysis would not have been feasible without the powerful built in software tools and the scripts he provided that enabled us to create concentric regions within granulomas.

Finally, I would like to thank the Gates Foundation for providing the funding that has made all of this research possible.

DEDICATION

I dedicate this thesis to my parents and my partner for supporting and encouraging me throughout this journey.

TABLE OF CONTENTS

Chapter 1: Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (Mtb), is a complex and dynamic disease resulting in a wide spectrum of disease forms. Control of TB mainly relies on effective chemotherapy but even with uncomplicated cases, treatment of drug-susceptible TB requires 6-9 months with a multidrug cocktail consisting of 2 months of rifampin, isoniazid, pyrazinamide and ethambutol followed by a continuation phase with rifampin and isoniazid for the remaining duration. Treatment of drug-resistant TB (MDR-TB) or extensively drug-resistant tuberculosis (XDR-TB) takes years and requires more complex multidrug regimens, that are associated with severe side effects and cure success is very low (<50%) (Chiang and Yew 2009).

Although TB is a systemic disease, its primary site of infection is the lungs. Pulmonary TB in humans can develop with a wide range in number of pulmonary lesions forming granulomas (1-2 or many) and with multiple histologically distinct granuloma types (Cronan 2022, Walter et al. 2023). Within the spectrum of granuloma types, one type, known as necrotic granuloma, is characterized by a rim of fibrosis and a central region of caseum containing a high burden of extracellular bacilli, while another type is characterized by cellular aggregations containing intracellular bacilli. The lesions also differ in their cell composition, fibrosis, and vascularization and altogether within the TB granulomas, the bacilli are found in different physicochemical and biological micro-environments. The diversity of the biological environments and the lack of vascularization within the granulomas influence penetration of antibiotics and contributes to phenotypical and metabolic changes of bacilli and development of multiple bacterial populations within the granuloma ultimately affecting drug effectiveness (Ramey et al. 2023, Ernest et al. 2021, Walter et al. 2023).

Peptides that specifically recognize and home to diseased tissues or that target the cause of disease and that can penetrate tissue barriers are being investigated for targeted therapeutic approaches, with some initial success for infectious diseases and cancer (Lehar et al. 2015); (Hussain et al. 2018); (Hussain et al. 2018). These approaches can also be used to identify peptides capable of serving as organ- and pathology-specific probes to monitor disease states, progression, and/or therapeutic effectiveness (Teesalu, Sugahara and Ruoslahti 2012). Advancing targeted drug therapies using homing peptides coupled to nanoparticles is currently being studied as a therapeutic approach to deliver drugs to the diseased tissues and organs (Hussain et al. 2018, King et al. 2016, Lingasamy et al. 2019, Lingasamy and Teesalu 2021). Silver particles (SNP) are of interest in this therapeutic approach (Locatelli et al. 2014, Willmore et al. 2016) because they are biocompatible, cytotoxic and have good drug loading capacity. Furthermore, SNP have enhanced fluorescence intensity due to plasmonic effect which facilitate tracking and biodistribution studies (Willmore et al. 2016).

A current limitation for advancing the targeted drug delivery using homing peptides-nanoparticles loaded with drugs is lack of methodologies to study biodistribution and penetration of the homing peptide-nanoparticles into diseased tissues and organs. *Ex vivo* colocalization with inflammation sites using imaging has been reported (Mann et al. 2016, Põšnograjeva et al. 2022). However, the experimental characterization and quantification of homing peptides have used semi-quantitative imaging analysis, in most instances of the fluorescence-tagged peptides, that are time-consuming and may be subject to bias (Bankhead et al. 2017). Recent advances in machine learning approaches are now being used for unbiased identification, characterization and quantification using *ex vivo* samples from preclinical and clinical studies (Bankhead et al. 2017); (Asay et al. 2020). In this regard, image analysis has benefited from QuPath analysis software, an open-source for bioimage analysis and digital pathology designed to be user-friendly. QuPath is a powerful batch-processing with scripting functionality that allows the user

to develop and share new algorithms to analyze whole slide images of complex tissues (Bankhead et al. 2017, Berben et al. 2020). A number of modified Qupath scripts for image analysis already exist (Berben et al. 2020; (Apaolaza, Petropoulou and Rodriguez-Calvo 2021), (https://qupath.github.io).

In this study, we used an existing peptide PL1 (PPRRGLIKLKTS) with known receptors (i.e., Fibronectin extra domain B (FN-EDB) and Tenascin C C isoform (TNC-C) identified initially for its ability to target different solid tumors, angiogenic sites, endometriosis lesions (Lingasamy et al. 2019, Tobi et al. 2021), to also show high affinity for Mtb-infected tissues as well (to be published elsewhere). Here we sought to develop an advanced algorithm platform modified from QuPath (https://qupath.github.io) to create semi-automated workflows to greatly increase the speed and throughput of whole slide image analysis of Mtb-infected lung tissues and granulomas. We created a protocol that quantifies the distribution of FAM-labelled synthetic peptides, and peptides coupled to fluorescence (Cy3) silver nanoparticles, and *ex vivo* immunofluorescence images stained with a receptor specific monoclonal antibody to assess spatial distribution, abundance, co-localization and tissue penetration *in vivo*.

Chapter 2: Materials and Methods

2.1 Mice

The digital images from pulmonary samples used for the development of the image analysis software package were derived from archived Balb/c and C3HeB/FeJ mouse studies conducted in the testing of a Phage library (to be published elsewhere). Briefly, female Balb/c and C3HeB/FeJ mice at 6-8 weeks of age were purchased from the Jackson Laboratories (Bar Harbor, ME). Mice were rested for at least one week prior to the studies. All protocols and use of these animals were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at Colorado State University (CSU). At the end point of the studies, mice were humanely euthanized by CO_2 narcosis. Lungs were collected from each mouse and processed for further analysis.

2.2 Infection

Animals were infected with a low dose aerosol (LDA) infection of *Mycobacterium tuberculosis* (Mtb; Erdman strain). A Glas-Col inhalation exposure system (Glas-Col, Terre Haute, IN) was utilized to infect the mice. The bacterial working stock and inoculum were prepared for a targeted infection of 100 CFU (Balb/c) or 50-75 CFU (C3HeB/FeJ) per mouse. The stock and inoculum were serially diluted and plated in triplicate onto 7H11/OADC agar plates followed by incubation 37C for 3 weeks when the colony forming units (CFU) appeared and were visible to the naked eye. To confirm bacterial deposition into the lungs, mice (n=5) were euthanized after the LDA infection, and their lung homogenates plated onto 7H11 agar plates followed by incubation at 37C for ~4 weeks. The remaining mice were rested for 4 weeks (Balb/c mice) or 12 weeks (C3HeB/FeJ mice) until they were used in studies explained below.

2.3 PL1-fluorescent tagged peptide

Homing peptide PL1 (PPRRGLIKLKTS) was tagged with Fluorescein amidites (FAM) as reported by (King et al. 2016). FAM is a dye intended for oligonucleotide synthesis designed for molecular assays, microscopy, that can be excited using at 493 nm light and visualized with 517 filter sets. FAM-tagged homing peptides (PL1-FAM) were administered via intravenous (IV; 0.1 ml) or intraperitoneal (IP; 0.3 ml) injection to Mtb infected C3HeB/FeJ mice. Groups of animals (n=3) were rested for 1 hour. Thereafter, the mice were euthanized, and full body perfusion was performed with 10 mL PBS followed by 10 mL 4% paraformaldehyde (PFA) in PBS. The five lobes from lungs of each mouse were collected and processed for imaging as explained below

2.4 Fluorescence Nanoparticle Preparation and administration to mice

Silver nanoparticles (SNP) were prepared as in (Willmore et al. 2016) and tagged with Cyanine3 (Cy3) dye, a bright, orange-fluorescent dye that can be excited using the 532 nm laser line and visualized with TRITC (tetramethylrhodamine) filter sets. The fluorescent SNPs were functionalized with PL1 (PL1- SNPs) or with biotin molecules (Ctrl-SNPs and used as negative controls;) as in (Tobi et al. 2021). Both SNP types were administered via intravenous (IV; 0.1 ml) or intraperitoneal (IP; 0.3 ml) injection to Mtb infected C3HeB/FeJ mice. Animals were rested for 1 hour and then euthanized; their lungs were prepared as in above studies for the PL1-FAM administration.

2.5 Immunohistochemistry

Selected lung sections were stained with an in house-made antibody for FN-EDB (PL1 receptor; University of Tartu) or P32 (LinnTT receptor; University of Tartu) conjugated with Alexa 647.

Immunostaining was done according to the Laboratory of Precision and Nanomedicine at the University of Tartu.

2.6 Processing of samples to generate the digital image dataset

The lungs were harvested and fixed in 4% PFA-PBS for 48 hours. The five lobes from each lung from each mouse were placed into a histology cassette and embedded in paraffin. The paraffin blocks were sectioned at 5 μm and mounted with Prolong Gold antifade reagent with DAPI (InvitroGen).

2.7 Image data acquisition

All slides were scanned at 40X magnification with resolution of 0.25 μm /pixel using multispectral automated PhenoImagerTM (Akoya Biosciences) equipped with an acquisition software Vectra Polaris 1.0 (Akoya). Imaging of the whole slides was done with the following filters; FAM-labeled peptides were selected using the opal 520 filter (ex 499 nm; em 525 nm), Cy3-SNP were selected with a filter with ex 530-560 nm; em 579-720 nm and Alexa 647 labeled monoclonal antibodies FN-EDB and P32 were selected with the Opal 690 filter (ex 676 nm; em 694 nm). All files of imaged whole slides were approximately 5GB in size and were saved as a *.qptiff.

2.8 Software

QuPath (https://qupath.github.io) is a Java application developed in Java 8, with a user interface written using JavaFX. Files of imaged slides in a *.qptiff format were imported into QuPath.

Code availability. Source code and documentation for QuPath modification are available at https://github.com/JPeezy00/Colocalization_Thesis

2.9 Statistical analysis. The statistical analysis was done with guidance from the Statistical Department at CSU. Data was input and graphed in Prism v 10.2 including the statistical analysis performed using Chi-square.

Chapter 3: Results

3.1 *In vivo* screening of a phage-peptide library and selection of the lung homing peptide

The schematic representation of the experimental design used to select lung homing peptides from a phage peptide library in Balb/c mice, with or without infection with Mtb, is shown in Figure 1. After *in vivo* administration of a phage library and amplification (to be published elsewhere), the best performing homing peptide was PL1 and therefore was selected for this study. The scramble LinTT1 peptide was chosen as negative control. For imaging studies, the homing PL1 peptide or biotin (as negative control) were tagged to fluorescence (Cy3) silver nanoparticle (PL1-SNP; Ctrl-SNPs respectively) (Fig. 1C) (Willmore et al. 2016). The PL-1 or LinTT1 synthetic peptides were tagged with fluorochrome molecules of FAM (PL1-FAM; LinTT1-FAM respectively) as previously reported (King et al. 2016).

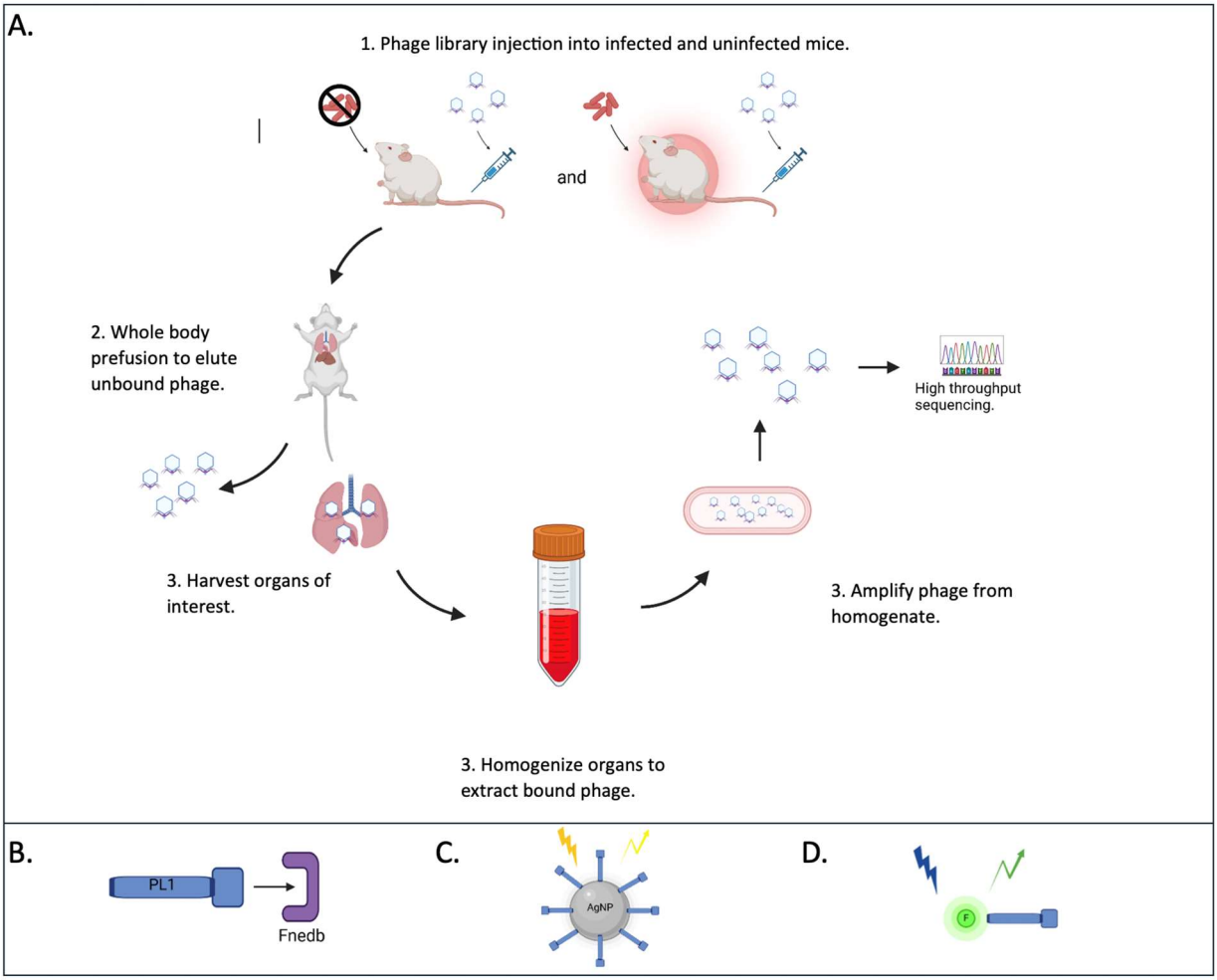

Figure 1. Phage selection and fluorescence functionalization. (A) Infected and uninfected mice were injected with a phage library presenting 22 binding proteins. After administration of the phage library, the mice were euthanized and perfused to elute unbound phages. The organs were harvested, the lungs were homogenized, and their phage content was amplified. Then, the amplified phages were analyzed through high throughput sequencing. (B) Protein PL1 was identified as the best cell-free performing homing to target FN-EDB. (C) Depiction of a Cy3 fluorescence silver nanoparticle with bound homing protein. (D) Depiction of a FAM fluorescence molecule with bound homing PL-1 protein.

3.2 Analysis strategy for identification and quantification of homing peptide within granulomas

C3HeB/FeJ mice infected with Mtb for 12 weeks received PL1-FAM, LinTT-scrambled-FAM, PL1-SNPs or Crt-SNPs via IV or IP injection as explained in Material and Methods. The whole slide imaging was performed using the Vectra Polaris (Akoya). The analysis was performed with modified Qupath scripts (available from https://github.com/JPeezy00 /Colocalization_Thesis) to create a protocol that quantifies the distribution of fluorescence in the lung sections. A flow chat of the QuPath protocol is shown in Figure 2. Briefly, the whole slide fluorescence image was used in the analysis (Fig 2A left). Ten annotations or regions of interest (ROI) were made for each class of tissue type corresponding to background (yellow squares), tissue (magenta squares), or granuloma (cyan squares) (Fig 2A right). Each ROI was manually selected by drawing a square with the "drawing" tool in QuPath (Fig 2A right). Next a multilayer perceptron (MLP) was trained to classify each ROI in the whole slide image. The ROIs were then created for each class shown in yellow (background), magenta (tissue) and cyan (granulomas) colors (Fig 2B). After training, ROIs were created for each thresholded class using "the create objects options" within the pixel classifier in QuPath. The background (yellow; Fig 2D) and stroma (magenta; Fig 2E) classes were then deleted leaving only ROIs corresponding to the lesion class (cyan color in red box; Fig 2F). Using a modified QuPath script, the ROIs corresponding to the granulomas were selected and divided into 3 concentric boundaries, outer, inner, and center (Fig 2 G insert and Fig 2 H, I, J respectively). All numerical data was extracted from QuPath and further analysis was performed using Prism or Microsoft Excel software.

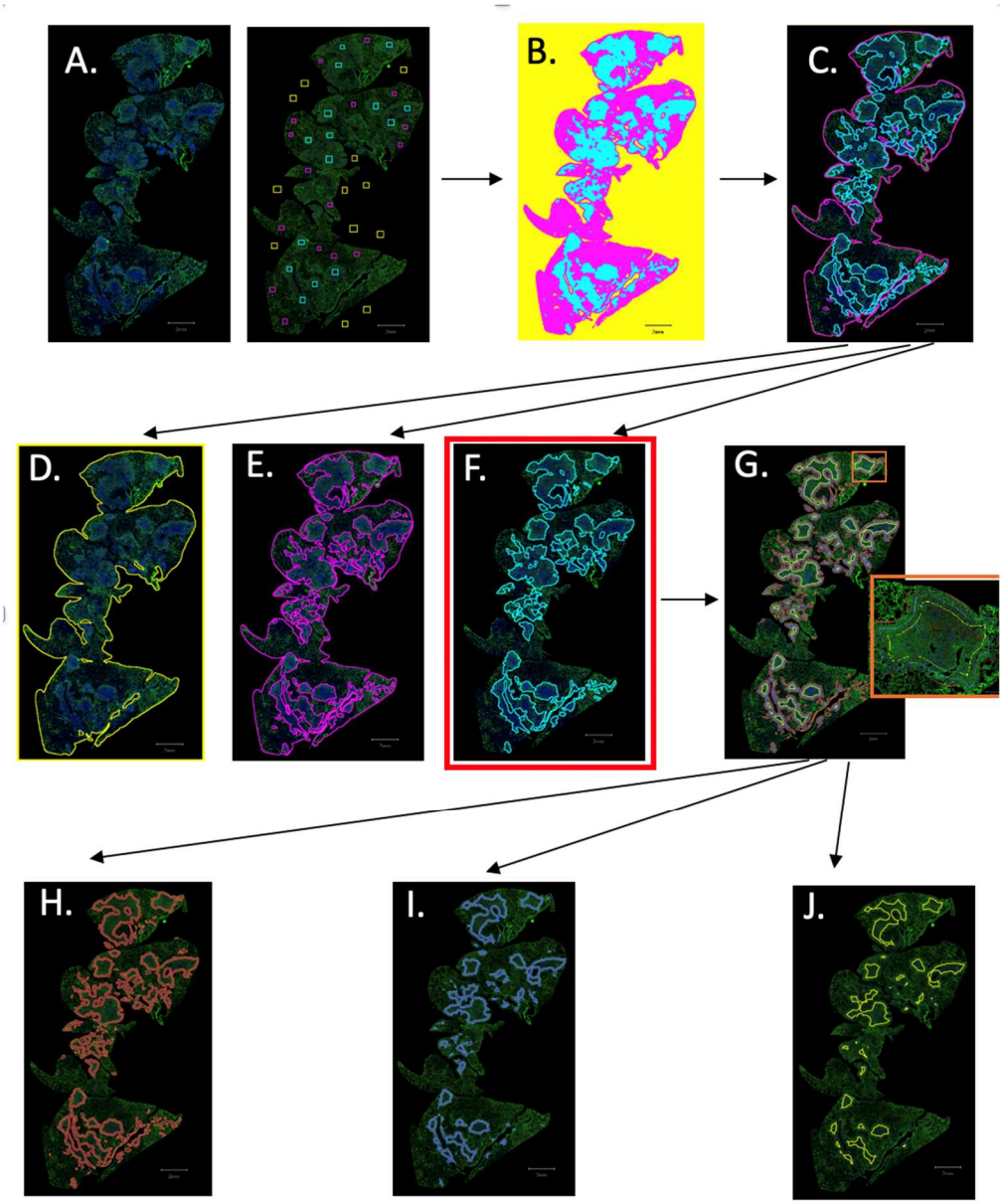

Figure 2. A flowchart for identifying granulomas and creating concentric regions within each granuloma using Qupath. (A) Whole slide images with affixed lung sections containing fluorescence SNPs or PL1 are digitized and uploaded to Qupath. At least ten annotations or region of interest (squares in A right) are manually selected as background (yellow squares), tissue (magenta squares) and granulomas (cyan squares) (B) Regions of interest (ROI) are used to train multilayer perceptron (MLP) to identify different classes (background, tissue, and granuloma) in the image; background (yellow), tissue (magenta), and granulomas (cyan). The trained MLP creates thresholded regions for each class (C-F). After training, the ROIs are created for each thresholded class using "the create objects options" within the pixel classifier in QuPath. The background and stroma classes are then deleted leaving only ROI for the granuloma class (red box). (G-J). Using modified Qupath script, additional concentric ROIs are drawn in G. Each concentric region is defined as outer (red boundary; H), inner (blue boundary; I) and center (yellow boundary; J) in the granulomas.

3.3 Comparing administration routes of PL1 and homing to the lung granulomas

The lung sections in whole slides fluorescence images from animals receiving PL1-FAM either by IP (Fig 3A) or IV (Fig 3B) routes or PL1- SNPs via IV (Fig 3C) were used to compare homing of PL1 peptide to the lung's granulomas. Using the analysis strategy shown in Fig 2, the outer (red), inner (blue) and center (yellow) regions of the granulomas were marked and the total percentage of PL1 positive cells after IP or IV administration were calculated. The total output of cells in the granulomas analyzed by QuPath detection tool was 572,590 cells for PL1-FAM via IP, 481,062 cells for PL1-FAM via IV, and 805,803 cells for PL1-SNPs via IV. This analysis demonstrated that after administration via IP, the percentage of cells positive for PL1-FAM within granulomas (76%) was higher than via IV (65%) ($p>0.0001$; Fig3D). Furthermore, a more in-depth analysis of the percentage of PL1-FAM positive cells within the outer, inner or center regions of the granulomas was also higher after IP than IV administration ($p>0.0001$) (Fig 3E). Importantly, a comparative analysis of PL1-FAM or PL1-SNP administrated via IV demonstrated significantly higher percentages of PL1-SNP positive cells (Cy3; 70%) than PL1-FAM (Fig 3F; 65%) and similar trends were found in the outer, inner regions of the granulomas (Fig 3G). In summary, the IP route provided higher levels of PL1 homing peptide to cells within granulomas and furthermore, because IP administration is technically easier than IV administration, we concluded that PL1 peptide, as a synthetic peptide or nanoparticle formulation, should be administered via IP. Moreover, our data here suggest also that homing to granulomas is higher by SNP than the PL1synthetic peptides, but this needs to be investigated further. However, we faced several limitations in this analysis. The comparative analysis of the administration route did not account for the differences in circulation times. In addition, we did account for the increase in total peptide coupled to SNP compared to FAM labeled peptide.

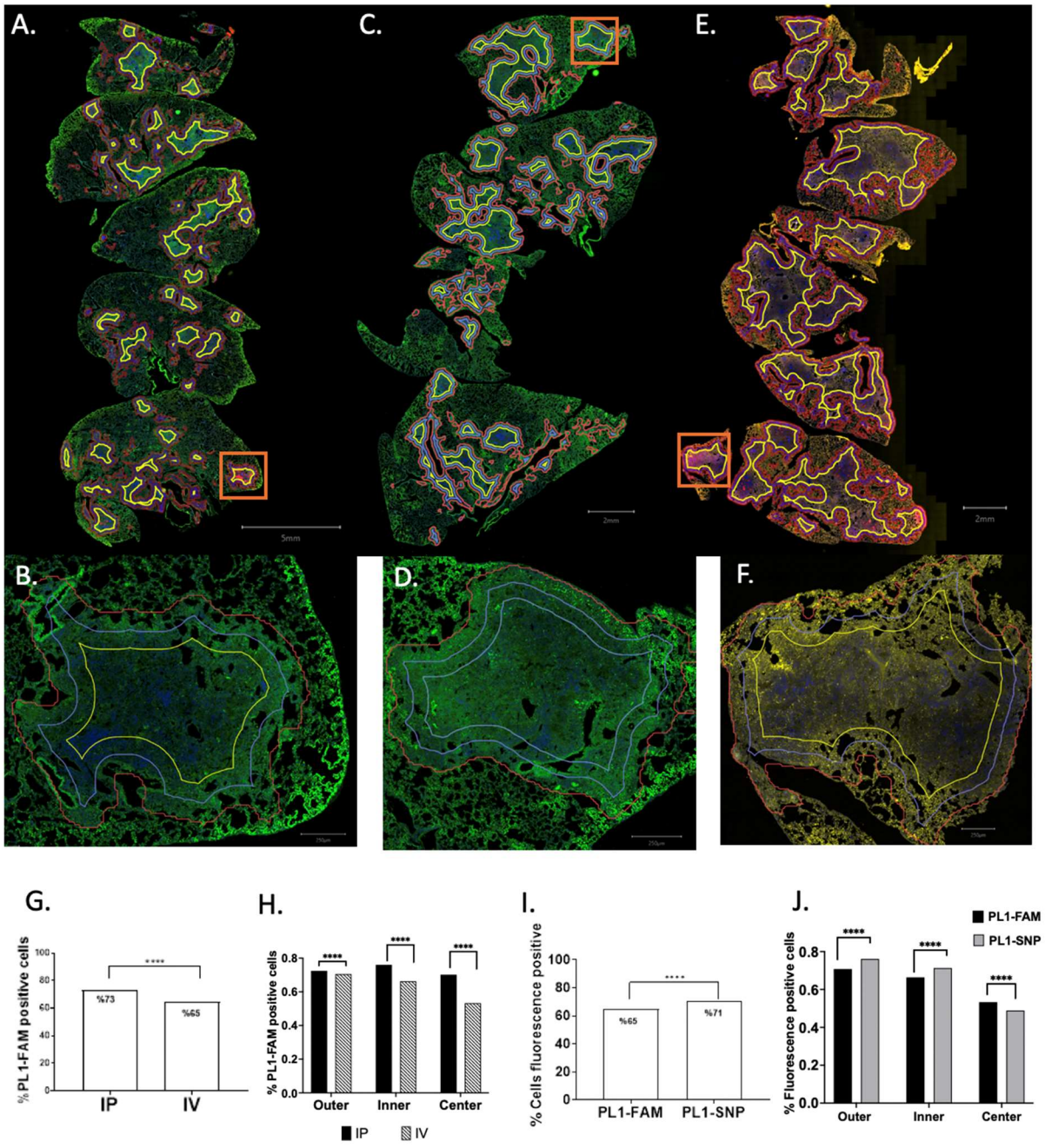

Figure 3.- Comparing administration routes of PL1 and homing to the lung granulomas. (A, B) PL1-FAM- injected intraperitoneally (IP). (C, D) PL1-FAM injected intravenously (IV). (E, F) PL1-SNP injected IV. The red, blue, and yellow outlines in A, B and C represent the outer, inner, and center margins of the granulomas, respectively. (G) Total percentage of PL1-FAM positive cells within granulomas of IP and IV injected mice. (H) Percentage of PL1-FAM positive cells within the outer, inner, or center concentric regions of the granulomas after IP or IV administration (I) Total percentage of PL1-FAM and PL1-SNP positive cells within granulomas and both administered IV. (J) Percentage of PL1-FAM and PL1-SNP positive cells within the outer, inner, or center concentric regions of the granulomas and both administered IV. **** $p<0.0001$

3.4 Quantification of PL1 within Lesions

In preliminary analysis, a conventional semiquantitative analysis using QuPath of whole slide fluorescence images showed colocalization of PL1-FAM with granuloma lesions whereas the scramble peptide LinTT1-scrambled lacked appreciable fluorescence signals in similar granulomatous lesions (Fig 4 A). To quantify the abundance of the PL1 peptide homing to the granulomas, the lung histology of a PL1-FAM dosed mouse was compared to a nonspecific scrambled LinnTT-scrambled–FAM (Fig 4) dosed mouse. To identify PL1 positive cells within granulomas, machine learning (multilayer perceptron or MLP) was used to accurately isolate different tissue types as explained in Fig 2. Then, the "positive cell detection" tool in Qupath was used to identify individual fluorescence-positive cells within the granulomas. The total output of cells in the granulomas analyzed by QuPath "detection tool" was 504,315 cells for PL1-FAM, for 858,084 cells for LinnTT-scrambled-FAM (). In Fig 4, cells positive for FAM can be seen in green while negative cells are shown in orange (Fig 4 A, B, E, F). Moreover, density heatmaps were created to help with visualization of concentration gradients for PL1 or LinnTT-scrambled within granulomas (Fig 4 C, D, G, H). The total percentage of PL1 (65%) and LinnTT-scrambled (40%) positive cells within the granulomas was extracted from QuPath and further analyzed using Prism software (Fig 4I). The results demonstrated that the percentage of cells positive for the PL1 peptide (FAM positive) were significantly ($p<0.0001$) higher than the scrambled LinnTT-scrambled-FAM positive cells.

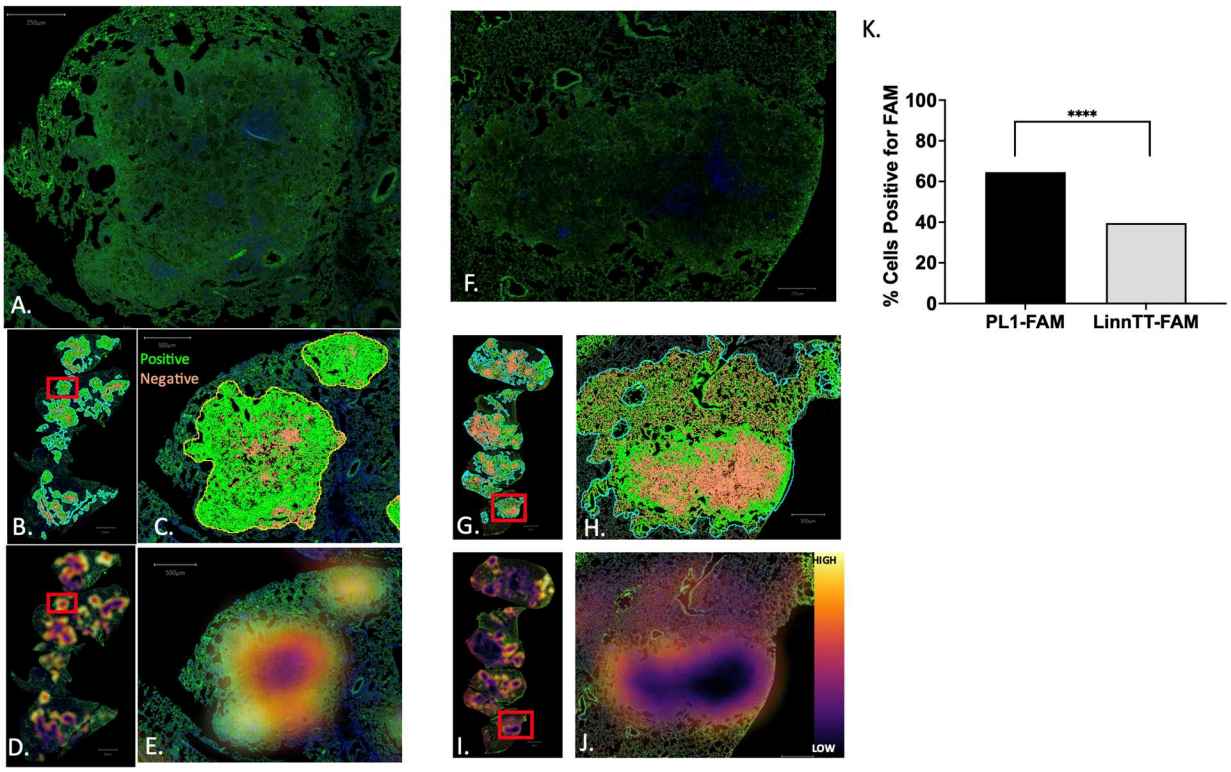

Figure 4. Quantification of PL1 within granulomas. (A, B) Granulomas within the lungs of mice injected with PL1-FAM or LinnTT-scrambled-FAM respectively. Using Qupath's "positive cell detection" tool with a threshold of 50; (B, C) PL1-FAM positive (green) and PL1-FAM negative (orange) or (G, H) LinnTT-scrambled-FAM positive (green) and LinnTT-scrambled-FAM negative (orange) cells were discriminated amongst. (D, E) Density maps created by Qupath of PL1-FAM positive cells (warmer colors) and (I, J) Lin-TT-scrambled-FAM cells (warmer colors) within granulomas (K) The total percentage of PL1 and Lin-TT Scrambled positive cells within whole granulomas. **** $p<0.0001$

3.5 Penetration of PL1 peptide within outer, inner, and center regions of the granulomas

To investigate the capacity of PL1 peptides to penetrate the granulomas (outer, inner, center), a modified QuPath script was used to separate ROIs within the granulomas into three concentric regions (Fig 5 A, B). The "built in" cell detection tool in QuPath was used to identify and quantify individual cells within each region of the granulomas (Fig 5 C-E). The total output of cells in each granuloma region analyzed by QuPath "detection tool" was 219,071 cells in the outer, 117,422 cells in the inner and 144,569 cells in the center. A normalized histogram for each marker (PL1 and LinnTT-scrambled) was created for each granuloma region (Fig 5F and G respectively). The graphs show the mean fluorescence bins on the x axis (equivalent to the intensity of fluorescence or abundance of PL1 per cell) and the total numbers of cells within each bin on the y-axis (Fig 5 F-H). The overall analysis in Fig 5H demonstrated firstly; the fluorescence of PL1 positive cells within the whole granuloma shifts right when compared to the LinnTT-scrambled fluorescence histogram, indicating higher levels of the fluorescence (and abundance) of PL1 per cell when compared similar LinnTT-scrambled. Secondly, there was a gradient in abundance of PL1 peptide homing from outer and inner to center of the granulomas, thus the highest numbers of high fluorescence of PL1 expression in inner followed by the outer and the center.

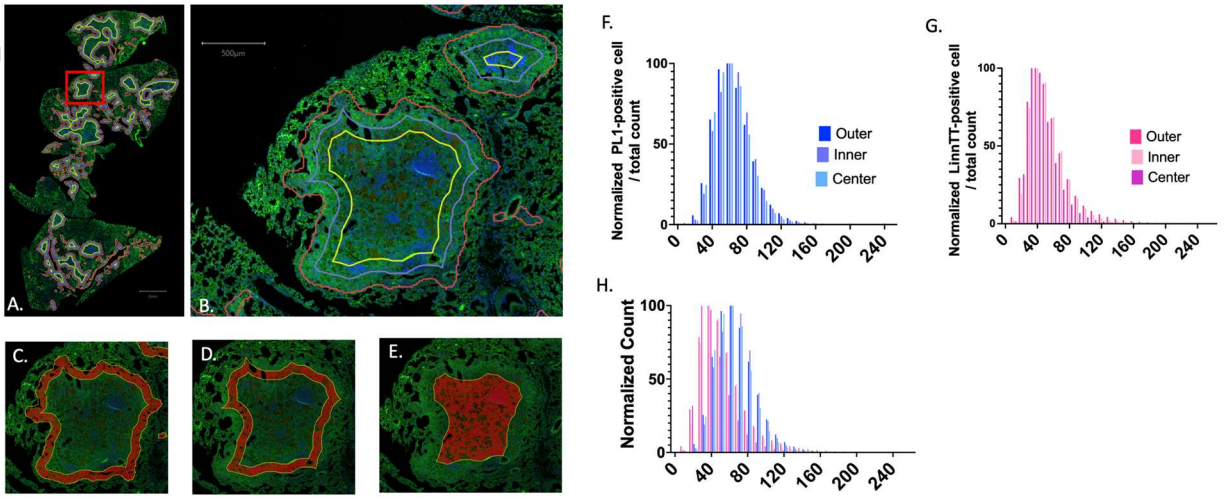

Figure 5. Distribution of PL1 positive cells within the granulomas. (A, B) A modified QuPath script was used to divide the granulomas into outer (red), inner (blue), and center (yellow) ROIs. (C-E) The QuPath's "cell detection" tool was used to identify individual cells (red) and measure the mean fluorescence for each cell in each region (outer, inner, and center) respectively. (F) Normalized histogram of the mean fluorescence intensity of PL1-FAM positive cells within each granuloma region; outer (dark blue), inner (blue), and center (light blue). (G) Normalized histogram of the mean fluorescence intensity of Linn-TT-FAM Scramble positive cells within each lesion region; outer (dark pink), inner (pink), and center (light pink). (H) Merged graph of F and G showing a swift to the right (high intensity of fluorescence) for the PL1-FAM positive cells.

3.6 Colocalization of PL1-FAM with its receptor FN-EDB compared with colocalization of LinnTT-scrambled-FAM and its receptor, p32

The results of the spatial colocalization within granulomas of the PL1 and LinnTT-scrambled with their corresponding receptors FN-EDB and P32 respectively are shown in Fig 6. Slides containing lung sections from animals receiving the PL1-FAM or LinnTT-scrambled-FAM followed by immunohistochemistry staining with fluorescence (Alexa 647) monoclonal antibodies against the FN-EDB or p32 respectively are shown in Fig. 6. Colocalization (green) or lack of colocalization (red) of both fluorescent signals (FAM and Alexa 647) for lung sections from mice receiving PL1-FAM (Fig 6 A, B, C and D) or from mice receiving LinnTT-scrambled-FAM are shown in Fig 6 E, F, G and H. Furthermore, additional concentric regions within granulomas were selected and marked as red (outer region), blue (inner region) and yellow (center region). The graph in Fig 6I shows the quantitative analysis of positive cells for the FN-EDB (black bars) and P32 (grey bars) within the outer, inner center region of the granulomas. This analysis demonstrated that both the PL1-FAM and LinnTT-scrambled-FAM treated mice show similar percentage of positive FN-EDB and P32 receptor in each granuloma region. The graphs in Fig 7 J show the quantitative analysis of positive colocalization for the FN-EDB with PL-1 and LinnTT-scrambled with P32 in the outer, inner, and center regions of the granulomas. This analysis demonstrated that the FAM-PL1 treated mice show higher percentage of positive colocalization when compared to the LinnTT-scrambled-FAM treated mice ($p<0.0001$) and therefore these analyses demonstrated that PL1, instead of LinnTT-scrambled, is the specific homing peptide in the lesions and within the outer, inner, or center regions of the granuloma.

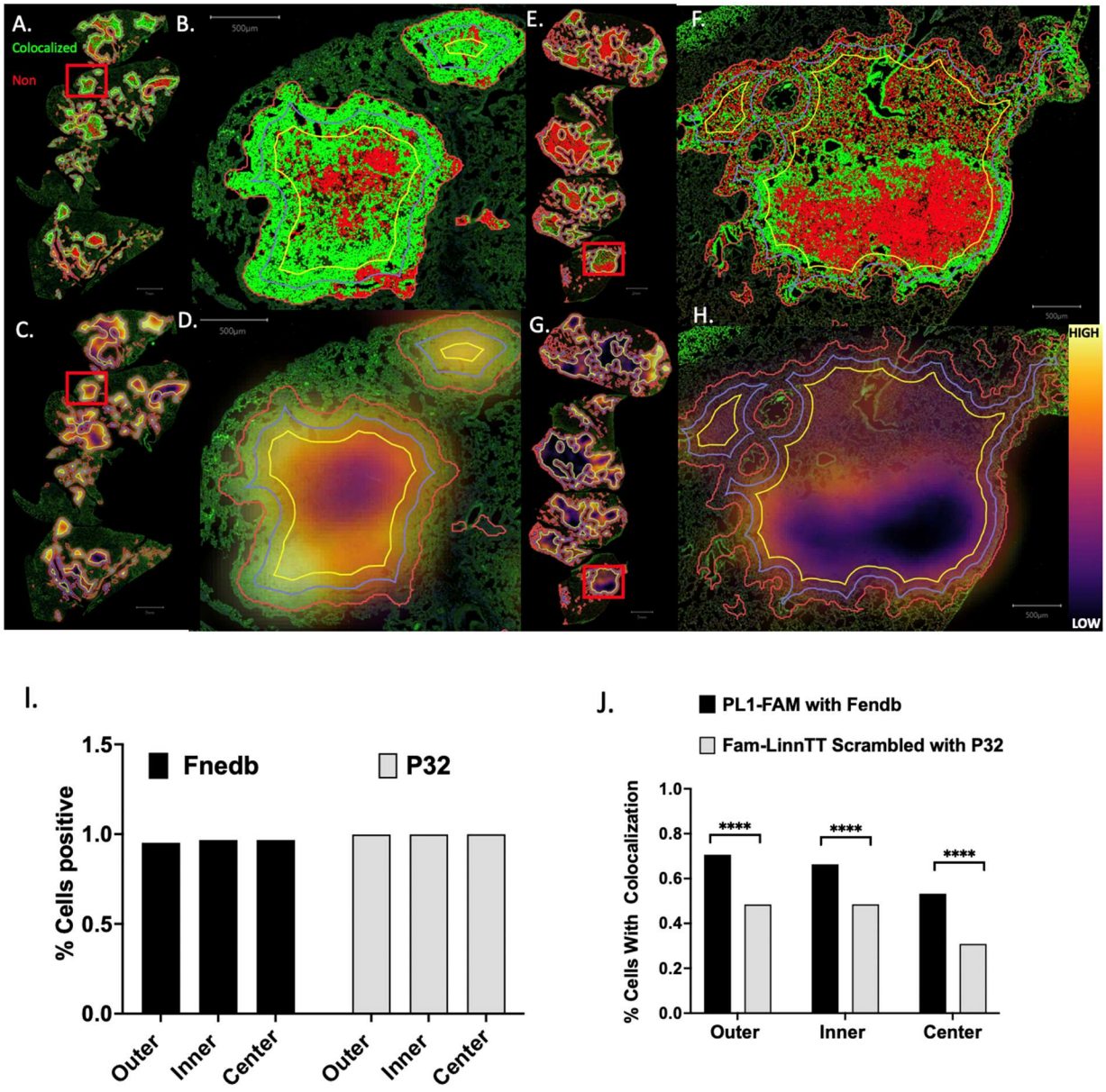

Figure 6. Colocalization of PL1-FAM with its receptor FN-EDB compared with LinnTT-scramble-FAM and its receptor, p32. (A, B) Cells colocalizing with PL1-FAM and anti FN-EDB – Alexa 647 (green) and cells with only one fluorophore or none (red). (C, D, G, H) A heatmap representing areas with colocalization (warmer colors), and low colocalization (cooler colors). (E, F) Cells with both LinnTT-scrambled-FAM and anti P32– Alexa 647 (green) and cells with only one fluorophore or none (red). (I) Percentage of cells within concentric regions (outer, inner, center) of the granulomas for FN-EDB black bars), or p32 (grey bars). (J) Percentage of cells within concentric regions (outer, inner, center) of the granulomas showing colocalization for both PL1and FN-EDB (black bars) and LinnTT-scrambled-FAM with P32 (grey bars). **** $p<0.0001$

3.7 Colocalization within granulomas of PL1-SNP with its receptor FN-EDB

The main goal of our studies is to deliver nanocarriers (SNP) containing drugs to be delivered into the granulomas and more important into deeper central regions of the granuloma. Thus, it was important to determine if the PL1-SNP also home to the granulomas and if they colocalize its intended target receptor FN-EDB (as in figure 6 with the synthetic homing peptide). A well accepted measure for correlation of target with its receptor is the Pearson's correlation analysis (Dunn, Kamocka and McDonald 2011). For this purpose, the lung sections from animals receiving the PL1-SNPs or Ctrl-SNP were also stained with a fluorescence (Alexa 647) monoclonal antibody against FN-EDB (Fig 7). The sections were then imaged and analyzed in QuPath as described above. The entire set of granulomas (ROIs) for both PL1-SNP (Fig 7A) and Ctrl-SNPs (biotin) (Fig 7B) were down sampled by a factor of 10 and extracted from QuPath using a modified script. After importing the images to image J (Schindelin et al. 2012) software, the channels were separated and "the colocalization threshold" tool (https://github.com/fiji/Colocalisation_Analysis?tab=readme-ov-file) in the Image J software was used to create a correlation scatter plot for PL1-SNP (Fig 7C) and Ctrl-SNPs (Fig. 7D) followed by determination of the Pearson's Coefficient for PL1-SNP and Ctrl-SNP (Fig 7E) along with the Average Pearson's coefficient of 41 random shifts in the X-axis of the channel 2 (anti FN-EDB-Alexa 647) image (Fig 7F). The overall results from these analyses demonstrated that there is higher colocalization of the PL1 with its receptor FN-EDB than with the Ctrl-SNPs and therefore it can be concluded that PL1-SNP has specific binding to its receptor.

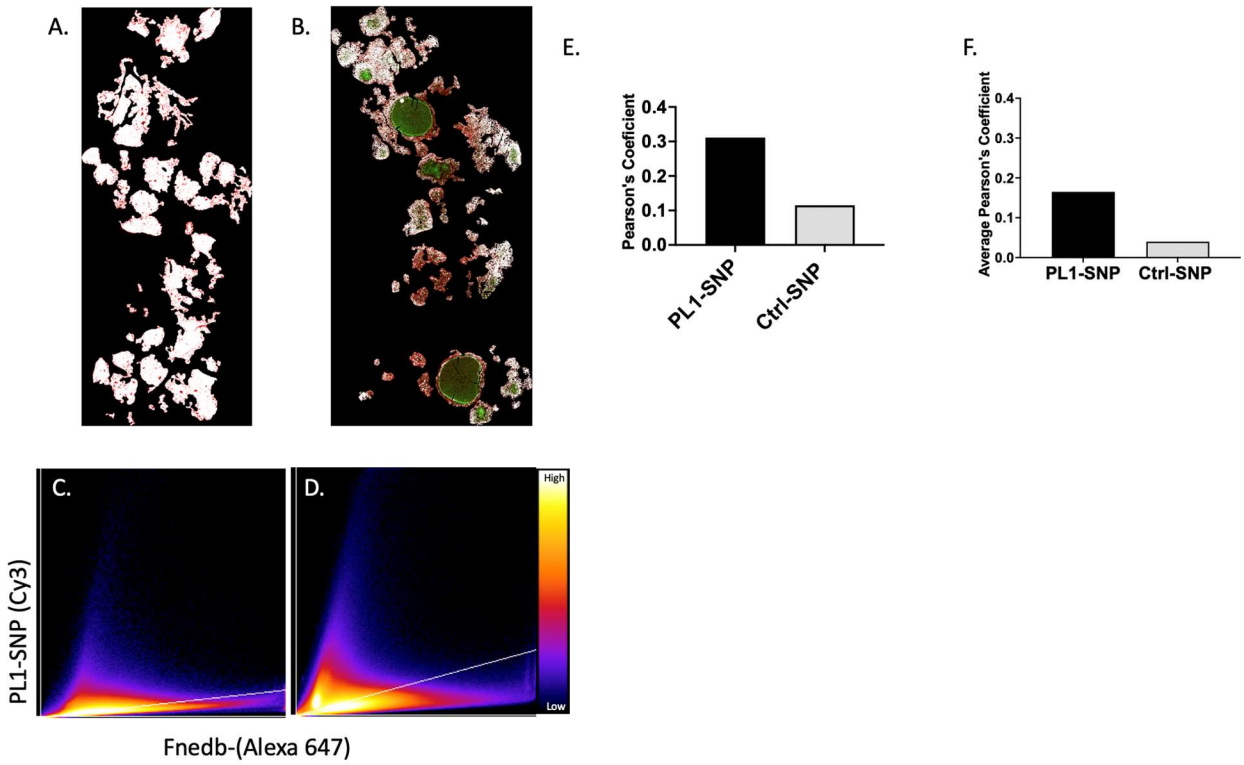

Figure 7. Pearson's correlation for colocalization of PL1-SNP with its receptor FN-EDB within granulomas. (A, B) Isolated granulomas (ROI) from whole fluorescence slide images of mouse lungs treated with PL1-SNP (Cy3) or Control-SNPs (Cy3), respectively. The slides with lungs sections in A and B were also stained with anti FN-EDB – Alexa-647 monoclonal antibody. The white regions represent areas of colocalization for SNP and FN-EDB, while green and red regions represent areas where either just FN-EDB only or SNP are present. (C, D) Pixel map of PL1-SNP (Cy3) (C) and Ctrl-SNP (Cy3) (D), respectively versus FN-EDB (Alexa-647). (E) Pearson's Coefficient of isolated granulomas. (F) Average Pearson's coefficient of 41 random shifts in the X-axis of the channel 2 (anti FN-EDB-Alexa 647) image.

3.8 Penetration of PL1-SNP within outer, inner and center regions of the granulomas

It was important to assess penetration and colocalization of the PL1- SNP within regions of the granuloma and association with its receptor FN-EDB. Thus, we analyzed the colocalization and created heatmaps of PL1- SNP (Fig 8A, B) or the negative control (Ctrl-SNPs) (Fig E, F) with the receptor FN-EDB within outer, inner, and center regions of the granulomas. Fig 8 shows representative images colocalizing (green) the PL1-SNP with FN-EDB (Fig 8 A, B, C, D) and Ctrl-SNP with FN-EDB (Fig 8E, F, G, H). Their corresponding heatmaps images are shown in Fig 8C, D and G-H. Numerical data extracted from QuPath from this analysis demonstrate that the abundance of the FN-EDB receptor (Alexa 647) (as in Fig 6I above) within each region of the lesions (outer, inner, and center) was similar (data not shown). Finally, the percentage of cells with colocalization of PL1-SNP and FN-EDB (Fig 8; black bars); Ctrl-SNP and FN-EDB (grey bars) demonstrate that PL1-SNP shows significantly ($p<0.0001$) higher percentage of cells than Ctrl-SNPs within the inner and center regions of the granulomas. It was interesting to see that the Ctrl-SNP's have poorer penetration or homing capability than the PL1-SNP into the inner and center regions of the granuloma. In other words, there is a much higher homing capability of the PL1-SNPs into deeper regions (inner and center) of the granuloma than the Ctrl-SNPs.

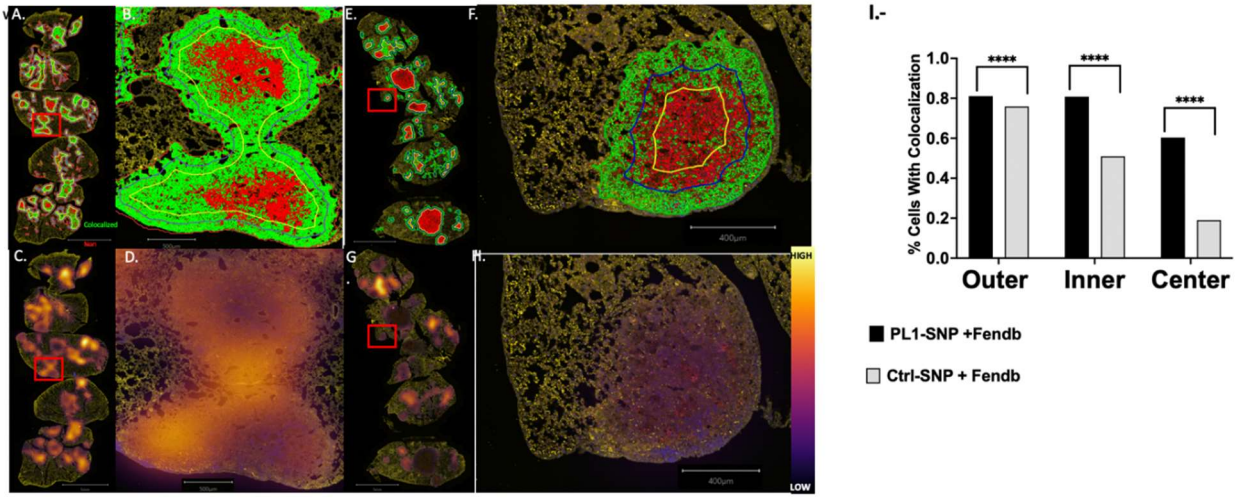

Figure 8. Assessing the specificity of PL1 binding with its receptor FN-EDB Colocalization of PL1-SNP (A, B) or the negative (biotin) control (Ctrl-SNPs) (E, F) with its receptor FN-EDB. (A, B) Cells colocalizing for PL1-SNP and FN-EDB (green) and cells with only one fluorophore or none (red). (E, F) Cells colocalizing for Ctrl-SNP with FN-EDB (green) and cells with only one fluorophore or less (red); (C, D, G, H) Heatmaps representing areas with colocalization (warmer colors), and low colocalization (cooler colors). (I) Percentage of cells within concentric regions (outer, inner, center) of the granulomas colocalizing PL1-SNP and FN-EDB (black bars); Ctrl-SNP and FN-EDB (grey bars). **** $p<0.0001$

Chapter 4: Discussion

Pulmonary TB is difficult to treat with chemotherapy and many TB drugs have poor penetration, or reach insufficient therapeutic levels into the granulomatous lesions where many extra and intercellular bacilli are located (Dartois and Dick 2024). We hypothesize that targeted delivery of TB drugs to the lungs allows for maximum efficiency of drug delivery, reduce systemic drug exposure and consequently also reduce adverse side effects. In cancer, chronic inflammatory and other infectious diseases, the use of homing peptides tagged to drug-loaded-nanocarriers holds great promises for therapy improvement (Pleiko et al. 2021). Targeted delivery can be achieved by identification first of homing peptides to diseased tissue (an unexplored arena in the TB field) and tagging of homing peptides to drug-loaded-nanocarriers (Pleiko et al. 2021). One limiting factor for development of targeted therapies is our limited capacity for unbiased colocalization and quantification of nanocarriers reaching the diseased tissue and their penetration capacity within the lesions. Here we developed a workflow in QuPath that reduces the bias of manually tracking and quantifying homing peptides and nanocarriers to granulomas of Mtb-infected mice.

The last decade has experienced great advances in the technology for fluorescence image acquisition, processing and analysis. In the past, the large surface area of tissues affixed to slides precluded scanning in a whole slide image and it required acquisition of images as multiple tiles. This process often required semiautomated stitching of the tile's images into one large image which was time consuming and left numerous artifacts. Moreover, manual annotations of regions of interest (ROI), localization and quantification of unique fluorescence signals within each slide were very time consuming and analysis was only performed manually in a limited sample size of ROIs within a slide. Mainly driven by the need for

immunophenotyping of tumors and efforts and investments of the cancer field in technology for more rapid and efficient fluorescence slide scanners (e.g. Akoya, Zeiss, Aperio), new software's for digital imaging analysis (Visiopharm, QuPAth, Inform) and introduction of machine learning tools, the spatial imaging analysis and phenotyping of cells within lesions have advanced significantly (Van Dam, Baars and Vercoulen 2022). Similar to immunophenotyping analysis of tumor microarrays (Van Dam et al. 2022), in our study, imaging and analysis of whole slides containing several lung lobes sections in the same slide provide more uniform and rapid analysis. Unlike commercial software (Visiopharm), scripts created by the developer of Qupath (Bankhead et al. 2017) have previously been modified by researchers (Berben et al. 2020) to accommodate new and very complex imaging analysis and their scripts have been made available in the GitHub as well as the Image Analysis forum. Altogether, the platform allows for rapid dissemination of open-source software for reevaluation within other types of imaging analysis. In our study, by using QuPath protocols, the time needed for tracking and quantification of granulomas and fluorescence particles in the whole section was dramatically improved. Moreover, application of machine learning tools (MLP) to isolate granulomas coupled with the analysis of approximately 400,000 – 1,000,00 cells is considerable shorter using the Qupath workflow developed here within and only limited by computer hardware capacity. All measurements for training of each type of tissue classifier (background, tissue, and granulomas) were performed automatically and were applied uniformly to other whole slide fluorescence images from the same or similar studies. Here we applied the same scripts and same machine training methodology to identify the tissue classifiers in all slides used in this study. It is also well established that the number of manual annotations made by the user in order to train the software how to identify tissue identifiers (background, tissue or granulomas) and the accuracy which this task is performed by the software have a large impact on classification by QuPath. In our study, we used 15 - 20 annotations for each tissue identifier and the same was validated in 6 whole slide images. A previously reported machine learning

imaging analysis, known as LIRA (Asay et al. 2020) identified and classified granulomas in the same animal models as in this study and the number of slides used to make annotations was 176 which were then split into over 1.5 million patches. Our modified QuPath and LIRA machine learning (Asay et al. 2020) both allow for identification of the granuloma lesions but while LIRA uses H&E stained slides, our trained QuPath granuloma classifier uses the fluorescence imaging and expands the analysis to tracking of fluorescence peptides and nanocarriers. The LIRA workflow allows for discrimination of necrotic versus non-necrotic granulomas, while our workflow here was not designed for such discrimination, mainly because the sample size for necrotic and non-necrotic granulomas was insufficient to perform such a task. Future modification of the script is planned to discriminate between several types of granulomas but this will require performing several additional studies to increase the sample size of granuloma types.

Penetration to TB drugs into the granuloma has been partly studied using MALDI (Ernest et al. 2021) and pharmacokinetics but it is a still a poorly understood area in the TB therapy field. The potential effectiveness of targeted therapies hinges on the ability of the drug carriers to infiltrate the granulomas. For this reason, the computerized QuPath protocol here was also modified to identify distinct regions in the granuloma (outer, inner, and center). The modified QuPath was trained to identify the outer, inner, and center regions and furthermore was able to enumerate the total number of cells in each region. Moreover, the QuPath script was also modified to discriminate and quantify abundance of fluorescence associated with PL1 (FAM) or SNP (CY3) and anti-FN-EDB antibody (Alexa 647). Altogether these tools within the modified QuPath script, allowed for colocalization of the PL1 within each granuloma region while also colocalized the PL1 with its intended target, the FN-EDB receptor. Moreover, it allowed us to determine the penetration capacity and specificity of binding of the synthetic peptide to cells within each region of the granuloma. Similar analysis using the fluorescence of Cy3 derived from the SNP allowed

to assess the penetration capacity of the nanoparticles into the granuloma, binding to cells within each region of the granuloma and specific recognition of the SNP for the PL1 receptor FN-EDB.

Specific association of PL1 peptide with cells within the outer, inner, and center regions of the granulomas were measured and showed a noticeable shift in fluorescence towards increased fluorescence levels in samples from the animals receiving the PL1-FAM versus LinnTT scramble-FAM treated animals. This data suggests a strong association of PL1 with cells in these regions. Importantly the level of fluorescence associated with PL1 was higher in the inner and center regions of the granulomas. Furthermore, based in the Pearson's correlation test, there was strong colocalization of fluorescence signals derived from PL1 and its receptor FN-EDB suggesting that PL1 peptide was specifically bound to its receptors and hence confirming homing to the peptide with not only the desired targeting organ, the lungs, but also to the center of the granulomas and to its receptor. Moreover, our data here demonstrate that while the abundance of receptors for FN-EDB (PL1) or the scramble LinnTT peptide (P32) in each region of the granuloma was similar, there was an increasing gradient of abundance of PL1 into the inner and center region of the granuloma. This data suggests that penetration of PL1 into the granuloma was higher than the control scrambled peptide and also targeted and directed towards its specific FN-EDB receptor. The results were further corroborated during the analysis for penetration capacity of SNPs into the granuloma. The PL1-SNP's were able to penetrate and associate with cells and their intended receptor within the inner and center core of the granulomas whereas the Control-SNP found a hard to penetrate barrier in the outer region of the granuloma. As mentioned above our study here was not able to discriminate between necrotic and non-necrotic granulomas which are vascularized. It is anticipated that penetration of drugs, peptides and nanocarriers into necrotic granulomas maybe limited when compared to non-necrotic granulomas. As we move forward in our project and the number of *ex vivo* lung

samples from additional studies are available, we will use the scripts here as templates to develop new tools capable of assessing differences in penetration of targeted therapy in several granuloma types. Moreover, our analysis here focused on analysis of PL1 homing to the lung only but additional studies using this QuPath workflow will study homing of the PL1 peptide to other organs mainly the liver and spleen.

REFERENCES

Apaolaza, P. S., P.-I. Petropoulou & T. Rodriguez-Calvo (2021) Whole-Slide Image Analysis of Human Pancreas Samples to Elucidate the Immunopathogenesis of Type 1 Diabetes Using the QuPath Software. *Frontiers in molecular biosciences.,* 8.

Asay, B. C., B. B. Edwards, J. Andrews, M. E. Ramey, J. D. Richard, B. K. Podell, J. F. M. Gutiérrez, C. B. Frank, F. Magunda, G. T. Robertson, M. Lyons, A. Ben-Hur & A. J. Lenaerts (2020) Digital Image Analysis of Heterogeneous Tuberculosis Pulmonary Pathology in Non-Clinical Animal Models using Deep Convolutional Neural Networks. *Scientific Reports,* 10.

Bankhead, P., M. B. Loughrey, J. A. Fernández, Y. Dombrowski, D. G. McArt, P. D. Dunne, S. McQuaid, R. T. Gray, L. J. Murray, H. G. Coleman, J. A. James, M. Salto-Tellez & P. W. Hamilton (2017) QuPath: Open source software for digital pathology image analysis. *Sci Rep,* 7**,** 16878.

Berben, L., H. Wildiers, L. Marcelis, A. Antoranz, F. Bosisio, S. Hatse & G. Floris (2020) Computerised scoring protocol for identification and quantification of different immune cell populations in breast tumour regions by the use of QuPath software. *Histopathology,* 77**,** 79-91.

Chiang, C. Y. & W. W. Yew (2009) Multidrug-resistant and extensively drug-resistant tuberculosis. *International Journal of Tuberculosis and Lung Disease,* 13**,** 304-311.

Dartois, V. & T. Dick (2024) Therapeutic developments for tuberculosis and nontuberculous mycobacterial lung disease. *Nature Reviews Drug Discovery*.

Dunn, K. W., M. M. Kamocka & J. H. McDonald (2011) A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol,* 300**,** C723-42.

Ernest, J. P., J. Sarathy, N. Wang, F. Kaya, M. D. Zimmerman, N. Strydom, H. Wang, M. Xie, M. Gengenbacher, L. E. Via, C. E. Barry, 3rd, C. L. Carter, R. M. Savic & V. Dartois (2021) Lesion Penetration and Activity Limit the Utility of Second-Line Injectable Agents in Pulmonary Tuberculosis. *Antimicrob Agents Chemother,* 65**,** e0050621.

Hussain, S., J. Joo, J. Kang, B. Kim, G. B. Braun, Z. G. She, D. Kim, A. P. Mann, T. Mölder, T. Teesalu, S. Carnazza, S. Guglielmino, M. J. Sailor & E. Ruoslahti (2018) Antibiotic-loaded nanoparticles targeted to the site of infection enhance antibacterial efficacy. *Nat Biomed Eng,* 2**,** 95-103.

King, A., C. Ndifon, S. Lui, K. Widdows, V. R. Kotamraju, L. Agemy, T. Teesalu, J. D. Glazier, F. Cellesi, N. Tirelli, J. D. Aplin, E. Ruoslahti & L. K. Harris (2016) Tumor-homing peptides as tools for targeted delivery of payloads to the placenta. *Science Advances,* 2**,** e1600349.

Lehar, S. M., T. Pillow, M. Xu, L. Staben, K. K. Kajihara, R. Vandlen, L. DePalatis, H. Raab, W. L. Hazenbos, J. H. Morisaki, J. Kim, S. Park, M. Darwish, B. C. Lee, H. Hernandez, K. M. Loyet, P. Lupardus, R. Fong, D. Yan, C. Chalouni, E. Luis, Y. Khalfin, E. Plise, J. Cheong, J. P. Lyssikatos, M. Strandh, K. Koefoed, P. S. Andersen, J. A. Flygare, M. Wah Tan, E. J. Brown & S. Mariathasan (2015) Novel antibody-antibiotic conjugate eliminates intracellular S. aureus. *Nature,* 527**,** 323-8.

Lingasamy, P. & T. Teesalu. 2021. Homing Peptides for Cancer Therapy. In *Bio-Nanomedicine for Cancer Therapy,* eds. F. Fontana & H. A. Santos, 29-48. Cham: Springer International Publishing Ag.

Lingasamy, P., A. Tobi, M. Haugas, H. Hunt, P. Paiste, T. Asser, T. Rätsep, V. R. Kotamraju, R. Bjerkvig & T. Teesalu (2019) Bi-specific tenascin-C and fibronectin targeted peptide for solid tumor delivery. *Biomaterials,* 219**,** 119373.

Locatelli, E., M. Naddaka, C. Uboldi, G. Loudos, E. Fragogeorgi, V. Molinari, A. Pucci, T. Tsotakos, D. Psimadas, J. Ponti & M. C. Franchini (2014) Targeted delivery of silver nanoparticles and alisertib: in vitro and in vivo synergistic effect against glioblastoma. *Nanomedicine,* 9**,** 839-849.

Mann, A. P., P. Scodeller, S. Hussain, J. Joo, E. Kwon, G. B. Braun, T. Mölder, Z. G. She, V. R. Kotamraju, B. Ranscht, S. Krajewski, T. Teesalu, S. Bhatia, M. J. Sailor & E. Ruoslahti (2016) A peptide for targeted, systemic delivery of imaging and therapeutic compounds into acute brain injuries. *Nature Communications,* 7**,** 11.

Pleiko, K., K. Posnograjeva, M. Haugas, P. Paiste, A. Tobi, K. Kurm, U. Riekstina & T. Teesalu (2021) In vivo phage display: identification of organ-specific peptides using deep sequencing and differential profiling across tissues. *Nucleic Acids Research,* 49**,** 11.

Põšnograjeva, K., K. Pleiko, M. Haugas & T. Teesalu. 2022. New Tools for Streamlined In Vivo Homing Peptide Identification. In *Methods in Molecular Biology*, 385-412. Springer US.

Ramey, M. E., F. Kaya, A. A. Bauman, L. M. Massoudi, J. P. Sarathy, M. D. Zimmerman, D. W. L. Scott, A. M. Job, J. A. Miller-Dawson, B. K. Podell, M. A. Lyons, V. Dartois, A. J. Lenaerts & G. T. Robertson (2023) Drug distribution and efficacy of the DprE1 inhibitor BTZ-043 in the C3HeB/FeJ mouse tuberculosis model. *Antimicrob Agents Chemother,* 67**,** e0059723.

Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak & A. Cardona (2012) Fiji: an open-source platform for biological-image analysis. *Nature Methods,* 9**,** 676-682.

Teesalu, T., K. N. Sugahara & E. Ruoslahti (2012) Mapping of vascular ZIP codes by phage display. *Methods Enzymol,* 503**,** 35-56.

Tobi, A., A. M. A. Willmore, K. Kilk, V. Sidorenko, G. B. Braun, U. Soomets, K. N. Sugahara, E. Ruoslahti & T. Teesalu (2021) Silver Nanocarriers Targeted with a CendR Peptide Potentiate the Cytotoxic Activity of an Anticancer Drug. *Advanced Therapeutics,* 4**,** 11.

Van Dam, S., M. J. D. Baars & Y. Vercoulen (2022) Multiplex Tissue Imaging: Spatial Revelations in the Tumor Microenvironment. *Cancers,* 14**,** 3170.

Walter, N. D., J. P. Ernest, C. Dide-Agossou, A. A. Bauman, M. E. Ramey, K. Rossmassler, L. M. Massoudi, S. Pauly, R. Al Mubarak, M. I. Voskuil, F. Kaya, J. P. Sarathy, M. D. Zimmerman, V. Dartois, B. K. Podell, R. M. Savic & G. T. Robertson (2023) Lung microenvironments harbor Mycobacterium tuberculosis phenotypes with distinct treatment responses. *Antimicrob Agents Chemother,* 67**,** e0028423.

Willmore, A. M., L. Simón-Gracia, K. Toome, P. Paiste, V. R. Kotamraju, T. Mölder, K. N. Sugahara, E. Ruoslahti, G. B. Braun & T. Teesalu (2016) Targeted silver nanoparticles for ratiometric cell phenotyping. *Nanoscale,* 8**,** 9096-101.