

**Developing quantitative image analysis pipelines for scoring  
histological panoramic images: Testing Rab24 as a possible  
biomarker for cancer**

**Sadaf Fazeli**

**(523366, Sadaf.S.Fazeli@utu.fi)**



**Faculty of Medicine**

Institute of Biomedicine, Medisiina C, Turku, Finland

Master's Degree Programme in Biomedical Sciences, Biomedical Imaging track

**Supervised by:**

Eeva-Liisa Eskelinen, PhD

Professor of Medical Cell Biology, Institute of Biomedicine, UTU

Mauricio Ramm, MSc

Doctoral candidate, Institute of Biomedicine, UTU

**Specialization field:**

Light microscopy imaging

## Abstract

Biomarkers are highly essential to improve diagnosis, confirm the diseases' development, and monitor the treatment. Biomarker discovery requires analysis of a large quantity of data which is aided by computational tools. One of the methods widely used in search of new biomarkers is immunohistochemistry of tissue samples.

Numerous tools are available to detect different cell types in tissues in histological sections; still, the need for more advanced and quantitative analysis is growing. The most successful paradigms to meet these novel needs are using deep learning-based networks.

Rab24, an atypical member of the Rab protein family, plays a role in the late steps of endosomal degradation, in mitochondrial plasticity, and in the clearance of autolysosomes in basal autophagy. Rab24 has been connected to neurodegeneration and cancer. It has been shown to be overexpressed in hepatocellular carcinoma (HCC) and to enhance HCC's malignant phenotype. These findings together indicate that Rab24 might be a potential biomarker for cancer, and its modulation might be used as a strategy for cancer therapy.

This project was undertaken to investigate the expression of Rab24 in different types of human cancers. Rab24 was detected by immunohistochemical staining in cancer tissue samples embedded in paraffin. For the evaluation of expression levels, detailed image analysis pipelines were developed to combine an open-source software called QuPath with a deep learning network, StarDist, in order to setup a robust quantitative cell detection compatible with histological panoramic images.

Based on our current analysis, 5 cancer types, including angiosarcoma, stomach gastrointestinal stromal tumor (GIST), rectal neuroendocrine carcinoma (NEC), liposarcoma and fibrosarcoma were selected as potential candidates for further investigation.

**Key words:** Rab24, Cancer biomarker, Bioimage analysis, Deep learning, QuPath, StarDist, Biomarker discovery

## **Abbreviations**

CMA - Chaperone-mediated autophagy

DAB - Diaminobenzidine

GDP - Guanosine diphosphate

GDI - Guanine nucleotide dissociation inhibitor

GIST - Gastrointestinal stromal tumor

GTP - Guanosine triphosphate

HCC - Hepatocellular carcinoma

HRP - Horseradish peroxidase

IHC - Immunohistochemistry

NEC - Neuroendocrine carcinoma

NET - Neuroendocrine tumor

PBS - Phosphate buffer saline

ROI - Region of interest

TMA - Tissue microarray

TBS - Tris-buffered saline

TBST - Tween in Tris-buffered Saline

WSI - Whole slide image

## Table of contents

<b>1. Introduction</b> .....	1
<b>1.1 Biomarker discovery and data quantification</b> .....	1
1.1.1 Digital pathology and open-source tools in biomarker discovery.....	1
1.1.2 QuPath.....	2
1.1.3 Limitations of QuPath .....	3
1.1.4 QuPath combination with StarDist .....	4
<b>1.2 Rab proteins</b> .....	6
1.2.1 Rab24 .....	7
1.2.2 Autophagy .....	7
1.2.3 Endosomal degradative pathway .....	10
1.2.4 Mitochondrial plasticity.....	10
1.2.5 Rab24 as a potential biomarker for cancer .....	11
<b>2. Aim and hypothesis</b> .....	11
<b>3. Material and methods</b> .....	12
<b>3.1 Western blotting for Rab24 validation</b> .....	12
3.1.1 Sample preparation.....	12
3.1.2 Running the samples of the gel .....	12
3.1.3 Antibody staining of blots .....	13
3.1.4 Western blot image analysis .....	14
<b>3.2 Immunohistochemistry</b> .....	15
3.2.1 Histological quantitative image analysis .....	17
<b>4. Results</b> .....	32
<b>4.1 Rab24 antibody validation</b> .....	32
4.1.1 Western blot for antibody validation .....	32
<b>4.2 TMA quantification</b> .....	35
4.2.1 Classification of tissues based on staining intensity .....	36
<b>5. Discussion</b> .....	47
<b>6. Acknowledgement</b> .....	49
<b>7. References</b> .....	50

## **1. Introduction**

### ***1.1 Biomarker discovery and data quantification***

In the past, cancer diagnosis used to rely solely on the morphological analysis of tissues and cells in histological sections of the tumor. This type of morphological diagnosis requires many years of experience, it is not quantitative, and it is challenging when it comes to the early stages of malignant diseases. Every cancer develops due to changes at the molecular level. Therefore, studying the associated molecules is the key to determine cancer's type and stage precisely. Biomarkers are specific diagnostic/prognostic molecules and changes in their expression may indicate physiological and pathological conditions. Biomarkers are also helpful to ascertain the classical pathological diagnosis and to show the efficiency of the treatment. Many studies have been conducted to show the correlation of biomarkers with different diseases, as exemplified by the following publications. Flieswasser and colleagues (2019) showed that a biomarker called CD70 is commonly expressed in solid and hematological tumor entities. Another study revealed CDK6 as a potential biomarker for non-acute promyelocytic leukemia patients, a subtype of acute myeloid leukemia (non-APL AML) (Liu et al., 2021). Finally, Chen and colleagues (2017) showed that Rab24 upregulation facilitates the growth and metastasis of HCC cells.

Early diagnosis is essential for cancer treatment as it increases the patient's chance of recovery. Thus, biomarker discovery has gained much attention in recent years, and cancer therapies are also highly linked to the progression of this field.

#### **1.1.1 Digital pathology and open-source tools in biomarker discovery**

Discovering new markers requires analysis of large quantities of samples, and this is time-consuming or even impossible to be done manually; hence, computational methods need to be used. In particular this applies to analysis of histological samples. Despite the advances in digital pathology, the existing software may not be completely functional for addressing the needs of the new challenges, such as detecting cells that differ in size and shape simultaneously in one tissue section. These demands give rise to the need for more complex and advanced image analysis strategies. Open-source software can be one of the best solutions in addressing these issues as it can be developed further by multiple users

worldwide and be combined with machine learning methods. Besides, open-source products are beneficial for both programmers and users with no programming background, due to the active and supportive imaging community. Open resources have revolutionized digital pathology, and many findings in cancer biomarker discovery would not have been possible without the presence of this technology (Humphries et al., 2021).

### **1.1.2 QuPath**

QuPath (Bankhead et al., 2017), short term for quantitative pathology, is an extensible bioimage analysis software with an open-source solution that was initially created for histological analysis, but fluorescence microscopy images can be analyzed using this software as well. It is capable of dealing with the massive size of whole slide images (WSIs) and enables the user to quantify the biomarkers without cropping or lowering the size of images taken by panoramic scanners. Thereby, it is more convenient for digital pathology compared to e.g. the widely used ImageJ-Fiji software.

QuPath supports many image formats such as Digital Imaging and Communications in Medicine (DICOM), TIFF, and BIF. For reading MRXS format, often 3DHISTECH's Slide Converter needs to be installed. The flexible WSIs viewer in QuPath enables the user to navigate freely across the panoramic images. Besides, it allows annotating and classifying the tissue and tumor manually and automatically, alongside the staining intensity estimation. Positively stained cells can be detected, counted, and scored. QuPath further gives the possibility to assess tissue microarrays (TMAs) and exchange of data with ImageJ. Once an optimized workflow is created, it can be saved through the command history and run on a series of images. More advanced analysis is possible by batch-processing and scripting. Scripting unlocks a vast range of possibilities (Humphries et al., 2021; Bankhead et al., 2017; Welcome to QuPath! — QuPath 0.2.3 documentation).

Instructions regarding the analysis, installation, and useful tutorials for QuPath can be found in references (Welcome to QuPath! — QuPath 0.2.3 documentation; QuPath | Quantitative Pathology & Bioimage Analysis).

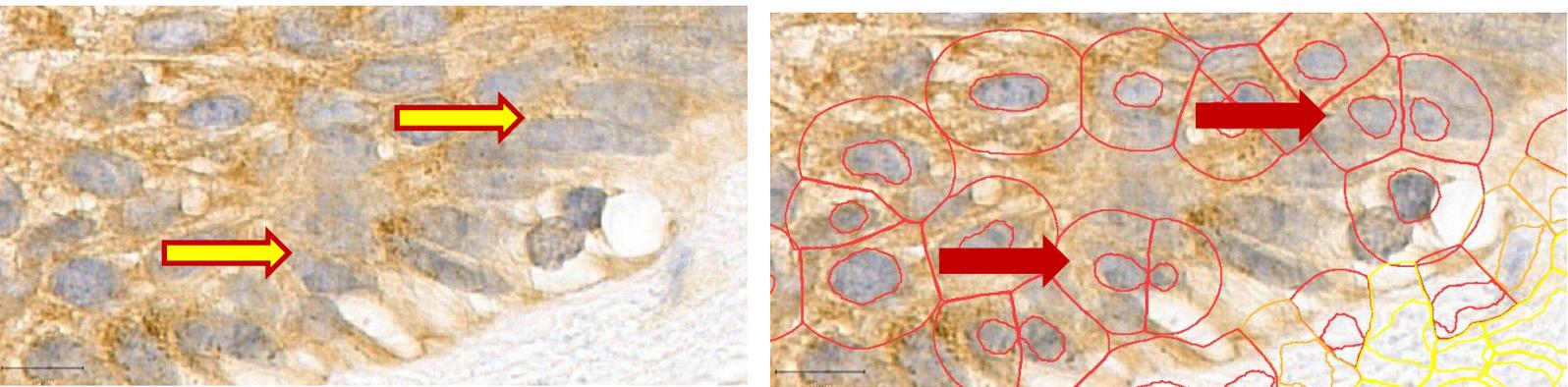
QuPath's supportive community website is designed for users with no programming background (Latest qupath topics - Image.sc Forum).

### 1.1.3 Limitations of QuPath

A reliable instant nuclei segmentation method allows analyzing microscopy images to identify cell shape, cell size, and gene and protein expression individually. This helps to answer biological questions such as gene expression differences between cells (Humphries et al., 2021).

QuPath is commonly used to detect, count, and score the cells from negative to 3+ (strong protein expression) in histologically stained tissues, and a list of commands in the software has been developed for this aim. The primary tool for cell detection and scoring in QuPath is the "positive cell detection" command.

The efficiency of the positive cell detection command depends on many factors, such as the intensity of positive staining, the number of overlapping nuclei, and the location of positive staining (nucleus or cytoplasm). It has been shown that in the images of confluent cells, or when the positive staining is localized in the cytoplasm, this command might not work accurately. This inaccuracy could result in nucleus over-segmentation, detection of false fragmented nuclei, and consequently unreliable and biased analysis. This error is shown in Figure 1.



**Figure1. False segmentation.** On the left, the original image, and on the right, the same image after applying positive cell detection in QuPath are shown. Yellow arrows indicate two nuclei. Red arrows correspond to the same cells in the right image which QuPath has incorrectly segmented into two, to demonstrate the over-segmentation error. Scale bar: 10  $\mu\text{m}$ .

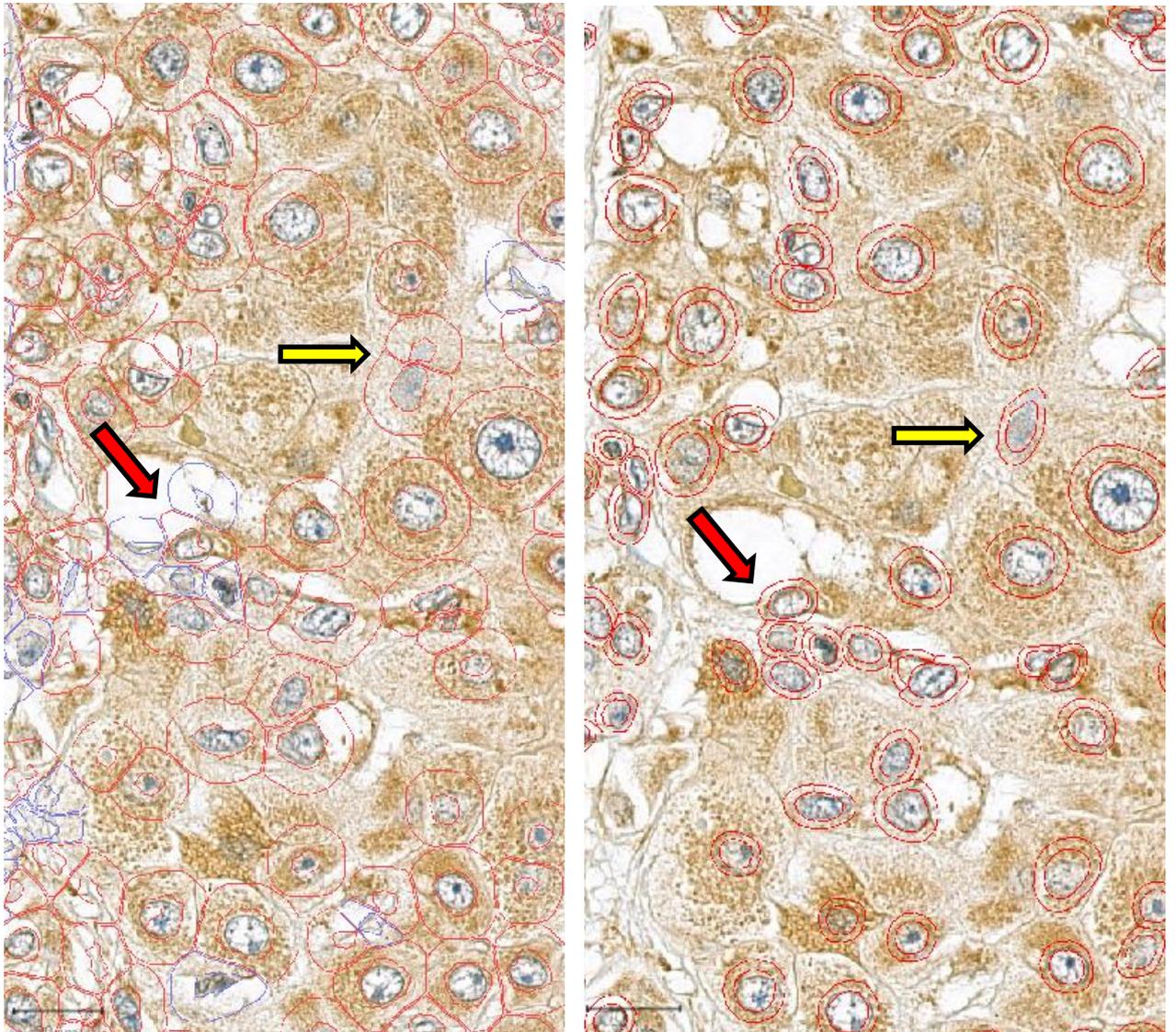
#### 1.1.4 QuPath combination with StarDist

The most successful paradigms to solve the nuclei detection issue are categorized in supervised deep learning. Using deep learning methods and their combination with current image analysis software enable the user to find all objects such as nuclei and delineate them in a very meticulous manner. In the deep learning method, the user formulates the task as a prediction test with deep learning neural networks. The input would be one or more images containing the structures of interest and the output would be a segmentation of the structures of interest. Firstly, user annotates a subset of the final dataset by segmenting the objects in the image manually or semi-automatically. These images are called the training dataset and are used to produce a model by the neural network. The model is then applied to the unannotated dataset in order to make segmentation predictions. Training must be optimized for every specific dataset in order to get the best result. Deep learning-based analyses are very promising for cell segmentation.

StarDist is a deep learning network for nuclei segmentation in 2D and 3D images, which is well adapted to nuclei's typical round shapes. It follows two steps; in the first step, instead of using foreground and background strategy (Bottom-up approach) for segmentation, StarDist predicts two parameters for each pixel. One parameter is an object probability that is computed from the ground truth, which is the distance of the center of each object to the next nearest boundary. This indicates the pixel's probability of being the center of the object. The second parameter is the radial distance, a certain number of radial directions that developers have predefined for every point. It indicates the distance to the closest boundary of the object. Therefore, it shows the distance of every pixel from the object boundary based on a predefined ray.

The following step is polygon selection and maximum suppression. Polygon is constructed out of the distance maps and object probabilities for every point and asks which point overlaps or which polygons assigned to a specific point have the highest probability. It also asks which points should be suppressed because the corresponding polygon overlaps too much with an already existing polygon, and this is called maximum suppression. The method finally keeps the most probable object according to the predicted probabilities and suppresses all the predicted polygons that overlap above a certain threshold. In Figure 2, cell detection, using "positive cell detection" command and

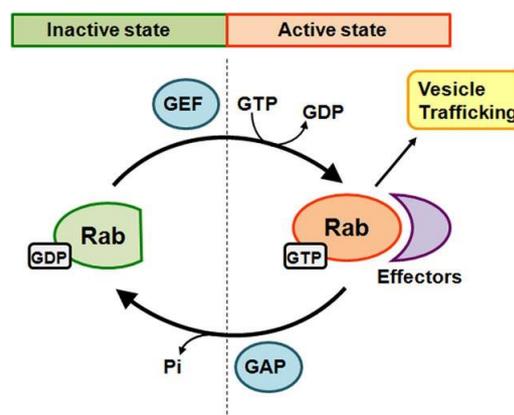
StarDist are compared with each other (Latest QuPath topics - Image.sc Forum; Schmidt et al., 2018; GitHub - maweigert/neubias\_academy\_stardist).



**Figure 2. Positive cell detection command in QuPath vs. StarDist.** The images show the same area analyzed by positive cell detection command (Left) and StarDist (right). The deep learning approach is better adapted to the nuclei shapes, allowing to simultaneously detect cells different in shape and size. It reduces the number of falsely segmented nuclei (Yellow arrows) and detects them more meticulously (Red arrows). Scale bar: 90  $\mu$ m.

## 1.2 Rab proteins

Rab proteins are soluble small GTPases that orchestrate all the steps in intracellular vesicular trafficking, such as cargo selection, vesicle budding and transportation, vesicle docking and fusion. Once Rab proteins are produced, they are placed in the cytoplasm for post-translational modifications, leading to the attachment of prenyl groups (hydrophobic lipid moieties) to their C-terminals. The prenyl group is inserted into the target membrane, enabling the Rabs to be anchored to intracellular membranes' cytosolic side. The localization of Rab proteins on the vesicular membrane is vital for their activation. GDP-bound cytosolic Rab (Rab-GDP) is the inactive form and, GTP- and membrane-bound is the active form of the members of this protein family (Figure 3) (Wang et al., 2017). The Rab protein is activated by the intervention of GDP-GTP exchange factors (GEFs). Active Rab can recruit its effectors and consequently perform a specific task. Inactivation of Rab occurs by the action of GTPase-activating proteins (GAPs), leading to hydrolysis of bound GTP to GDP and phosphate. In the deactivation process, Rab-GDP (inactive form) associates with GDP dissociation inhibitors (GDIs) and detaches from the membrane. Detachment is due to the GDIs hydrophobic interaction with the Rabs' prenyl group; thus, the prenyl group can no longer be inserted in the membrane. Dissociation of inactive Rab from GDI allows Rab to be activated again. As a result of this dissociation, Rab is able to insert its prenyl group in the target membrane, which occurs by the action of the GDI displacement factor (GDF) (Hutagalung and Novick, 2011). The Rab cycle is summarized in Figure 3.



**Figure 3. Inactive and active states of Rab proteins.** GTP- and membrane-bound is Rab's active form, while GDP-bound cytosolic Rab is inactive. The protein is recruiting its effector once it is in the active state. Inactivation occurs when bound GTP hydrolyzes, producing GDP and phosphate. The image is from reference (Wang et al., 2017).

Approximately 70 Rab proteins have been identified in humans, and each performs specific tasks based on the effectors it recruits. Several Rab proteins function in the autophagic process, including Rab1, 5, 7, 8B, 9A, 11, 23, 24, 25, 32, and 33B. Rab8B and Rab7 play an essential role in the maturation of autophagosomes, while the role of Rab8A and Rab25 is still unclear. Rab24 is needed for the clearance of autolysosomes (Ao et al., 2014; Ylä-Anttila and Eskelinen, 2018; Ylä-Anttila et al., 2015).

### **1.2.1 Rab24**

Rab24 is an atypical member of the Rab family, identified by Olkkonen et al. (1993). Rab24 has lower GTPase activity than the other members of the Rab family, which is likely due to serine instead of glutamine at position 67 in the GTP binding region (Ding et al., 2003). Different stimulators that trigger autophagic activity, such as stress conditions, have been shown to upregulate Rab24 expression in some cell types, which is in line with Rab24's role in autophagy. Eskelinen group (Ylä-Anttila et al., 2015) has revealed that Rab24 functions in the clearance of autolysosomes under basal conditions. Apart from autophagy, Rab24 plays a role in the late steps of the endosomal degradative pathway (Amaya et al., 2016), and additionally, it controls mitochondrial plasticity in the liver by interacting with the mitochondrial fission protein Fis1 (Seitz et al., 2019). The exact molecular mechanisms of Rab24 in autophagy are still unclear.

### **1.2.2 Autophagy**

Autophagy is a dynamic, highly regulated intracellular degradation mechanism, which takes place in eukaryotic cells. It breaks down redundant intracellular components within the lysosome to clean the cell of toxic proteins, defends against pathogens, and recycles damaged organelles. The degradation products can then be reused to produce energy, maintain homeostasis, and promote healthy cells' regeneration. Therefore, autophagy is a critical process for cellular homeostasis, and its dysregulation is connected to several diseases (Klionsky, 2008).

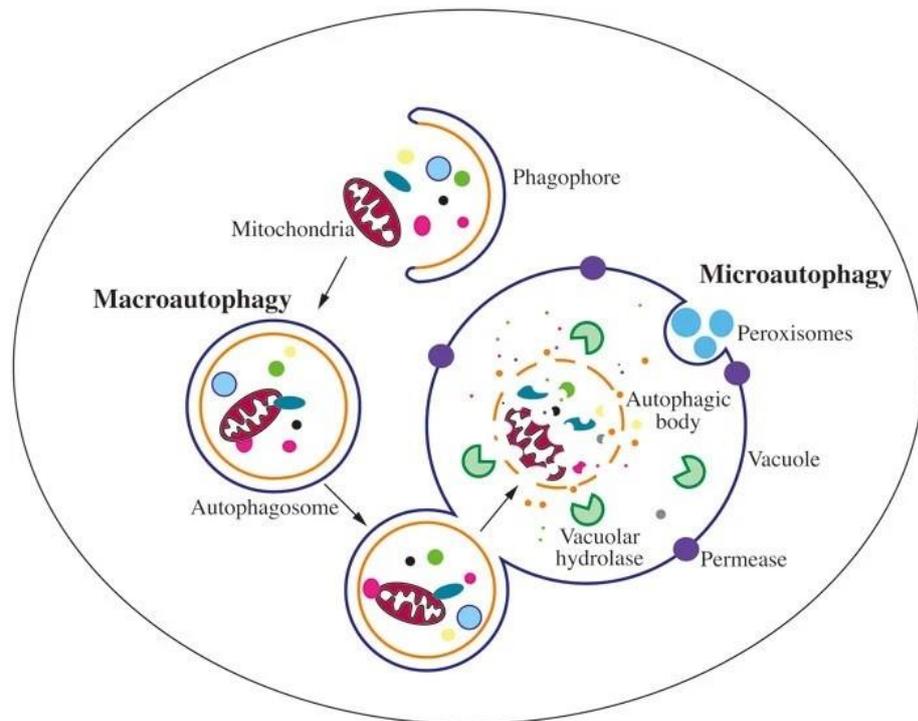
The general term autophagy refers to 3 types of autophagy: Macroautophagy, microautophagy, and chaperone-mediated autophagy (CMA) that all share the same characteristic, which is the degradation of cytosolic components within the lysosome.

In microautophagy, the degradation occurs after the direct engulfment and internalization of cytoplasmic components at the lysosomal membrane via invagination and formation of intralysosomal vesicles (Parzych and Klionsky, 2014).

CMA is a selective degradation process that targets soluble intracellular proteins with specific amino acid sequences, serving as "zip codes", exposed on their surface. The process requires chaperone proteins to recognize the zip codes, transport of the proteins to the lysosomal surface, and delivery of the cargo through the lysosomal membrane directly into the lysosomal lumen for degradation. CMA occurs during long periods of fasting, typically longer than 12 hours in cultured cells (Dong et al., 2020).

Macroautophagy is a highly regulated degradation process that begins at the endoplasmic reticulum. Cellular components are engulfed by a developing, cup-shaped membrane called the phagophore, leading to the formation of a double membrane autophagosome. The autophagosome transports the selected cargo to a lysosome-rich area, fuses with the lysosomal membrane, and releases the autophagic cargo inside the lysosome for degradation (Figure 4) (Feng et al., 2014). Compared to CMA, macroautophagy is induced by a shorter period of starvation (Mizushima et al., 2002).

Macroautophagy is an umbrella term for various types of autophagy that all utilize autophagosomes. Macroautophagy can be either non-selective or selective. Different types of selective autophagy are named according to the cargo, i.e., mitophagy, reticulophagy, nucleophagy, lipophagy, aggrephagy and xenophagy which refer to the degradation of mitochondria, endoplasmic reticulum, nucleus, lipids or lipid droplets, aggregates or protein condensates, and intracellular pathogens, respectively.



**Figure 4. Macroautophagy.** Selected cargo is engulfed by a cup-shaped structure called the phagophore that develops and leads to autophagosome formation. Autophagosome fuses with the lysosome, or vacuole in yeast, and releases the autophagic bodies inside the lysosomal or vacuolar lumen. The cargo is degraded by the lysosomal hydrolases and enzymes. The drawing also shows microautophagy of peroxisomes. Permeases transport degradation product back to the cytoplasm. The image is from Reference (Feng et al., 2014).

Autophagy is subdivided into two modes: Induced and basal autophagy. Induced autophagy is caused by stress such as starvation, hypoxia, the presence of pathogens, and chromosomal damage. On the other hand, basal autophagy constantly occurs at low level in most cell types under basal conditions when nutrients are present, to maintain homeostasis (Russell et al., 2014).

The autophagic process declines with aging, which causes impairment in cellular and tissue homeostasis. This reduction is partially due to the diminished proteolytic ability of lysosomes and the upregulation of the protein Rubicon, which inhibits autophagosome maturation. The less autophagic activity, the more toxic proteins accumulate, and dysfunctional organelles also become prone to accumulation. This accumulation can result in diseases such as neurodegeneration and cancer. Hence, autophagy is receiving much attention due to its potential role in longevity and preventing diseases such as cancer (Yamamuro et al., 2020). Autophagy may promote cell death and suppress the tumor

initiation in early tumorigenesis or, paradoxically, behave in favor of the cancer cell survival and metastasis in later stages of cancer (Yun and Lee, 2018). Autophagy modulators (both inhibitors and activators) are currently being used in clinical trials (Marinković et al., 2018).

### **1.2.3 Endosomal degradative pathway**

Extracellular components are internalized into cells through endocytosis. The cargo is transported in plasma membrane-derived vesicles to early endosomes (EE). Subsequently, they are either recycled back to the plasma membrane or transported to late endosomes (LE) for fusion with lysosomes for degradation. There are several regulators for the endocytic pathway, including some of the Rab protein family members. Rab 7 and Rab24 are localized in LE compartments. Rab7 regulates the late endocytic membrane trafficking and functions in the maturation of endosomes, similar to its role in autophagy. RILP (Rab interacting lysosomal protein) is a downstream effector that binds to Rab 7 in the active form. Rab24-Rab7-RILP complex allows endosome-lysosome fusion. Rab24 is also essential for activating Rab7 and for its attachment to the membrane (Amaya et al., 2016).

### **1.2.4 Mitochondrial plasticity**

Mitochondria are dynamic organelles that convert energy for the cellular processes due to the different nutritional conditions or energy status changes, to optimize ATP production in the cells. Mitochondria undergo dynamic fusion and fission. Rab24 was recently shown to regulate the mitochondrial fission in liver, by interacting directly with mitochondrial fission 1 protein (FIS1). This interaction is suggested to assist the recruitment of dynamin-related protein 1 (DRP1) to the mitochondrial membrane to drive fission. Mitochondrial fission plays a role in organelle plasticity under basal conditions. Fission is also required for mitochondrial degradation via mitophagy. Therefore, Rab24 inhibition in liver leads to less mitochondrial fission, less mitophagy, and more mitochondrial connectivity (Seitz et al., 2019).

### 1.2.5 Rab24 as a potential biomarker for cancer

Chen et al. (2017) showed that the over-expression of Rab24 facilitates the growth and metastasis of HCC cells both *in vitro* and *in vivo*. HCC is a lethal type of liver cancer, which occurs in patients with chronic liver diseases. The pathogenesis of HCC is not well understood. Recently, many reports have described a regulatory class of RNAs called microRNAs (miRNAs), which play a role in the pathogenesis of many cancers, including HCC. MicroRNA-615-5p was shown to be downregulated in HCC. This microRNA downregulates Rab24 expression; thus, the low level of miRNA-615-5p leads to the upregulation of Rab24, which was shown to lead to the promotion of cell migration, epithelial-mesenchymal transition, and ultimately invasiveness of HCC (Chen et al. 2017). The overexpression of Rab24 in HCC has been confirmed by two other research groups (He et al., 2002; Qin et al., 2021). In addition, Human Protein Atlas (Expression of RAB24 in cancer - Summary - The Human Protein Atlas) shows and high expression of Rab24 is a poor prognostic marker in liver and renal cancer. Cosmic cancer mutation database (COSMIC | Catalogue of Somatic Mutations in Cancer) shows that Rab24 is overexpressed in 14.5% (54 of 373) of HCC cases and in 11% (67 of 600) of renal cancers.

## 2. Aim and hypothesis

Despite all achievements in cancer biomarker discovery, for certain cancer types no biomarkers have been discovered yet. Biomarker insufficiency may lead to late diagnosis, cancer progression, and less chance of survival. Biomarkers are also used as tools to monitor the effectiveness of cancer treatments. Rab24 has been shown to be overexpressed in HCC and renal cancer, and we hypothesize it might be a potential biomarker for other cancer types as well.

This research project aimed to study the expression of Rab24 protein in 21 different human cancers using immunohistochemical staining of tissue microarrays obtained from the Helsinki biobank. To analyze the intensity of Rab24 staining, I developed quantitative image analysis pipelines using QuPath software combined with the deep learning network StarDist. The aim was to analyze histological images quantitatively and therefore, less subjectively.

### **3. Material and methods**

Tissue microarray (TMA) blocks containing samples from 21 different types of human cancer belonging to 54 patients (n=54) were obtained from the Helsinki biobank. The paraffin sections of these TMAs were stained for Rab24 using immunohistochemistry (IHC). Antibody validation for Rab24 antibodies was performed using Rab24 knockout, overexpressing and control cells in Western blotting. The following Neuro-2a cell lines were used: wild type (WT), Rab24 knockout (10G1), and GFP-tagged Rab24 overexpressing.

#### ***3.1 Western blotting for Rab24 antibody validation***

##### **3.1.1 Sample preparation**

In this section, all the steps were performed on ice.

The cells were scraped off the dish and collected in 1X PBS (140 mM NaCl, 2.7 mM KCl, 4.6 mM Na<sub>2</sub>HPO<sub>4</sub> 2H<sub>2</sub>O, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) followed by pelleting for 3 minutes at 400 g at 4°C. Next, the supernatant was aspirated, and the pellet was resuspended in lysis buffer (RIPA; 50 mM Tris-HCl pH7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 and Pierce protease inhibitor mini tablets). The suspension was then rotated for 40 minutes at +4°C. After the incubation, the suspension was centrifuged at 16000 g for 15 minutes at +4°C. Subsequently, the supernatant containing the total extract was collected and stored on ice. The total extract's protein concentration was measured using the BCA protein assay (Thermo Scientific, 23228). Absorbance was measured with Ensign Multimode Microplate Reader (Perkin Elmer) at 562 nm. Samples were prepared by adding 4X loading buffer (Laemmli buffer, 0.2 M Tris-HCl, pH 6.8, 8% sodium dodecyl sulfate (SDS), 20% 2-mercaptoethanol, 0.004% bromophenol blue), then vortexed and boiled at +95°C for 5 minutes. Ultimately, the samples were stored at -20°C until further use.

##### **3.1.2 Running the samples on the gel**

12% SDS-PAGE gel was prepared to run the samples. The samples were melted and boiled at +95°C for 3 minutes. 10 µg of total protein were loaded in each well alongside with 5 µl of molecular weight standard PageRuler Prestained Protein Ladder Plus (Thermo scientific) on the gel. The gel was run in MiniPROTEAN Tetra Cell apparatus

(Bio-Rad) filled with 1X running buffer initially at 50V for 20 minutes followed by 120V for 90 minutes. Next, the gel was incubated in 1X complete blotting buffer (Tris 0.025M, Glycine 0.2M, Methanol 20%, pH 7.6) on the shaker for 30 minutes. Afterwards, PVDF-membrane (Roth) activation was done in methanol for 20 seconds and then the membrane was equilibrated in blotting buffer for 15 minutes. Gel was placed on PVDF and run at 250 mA for 90 minutes at +4°C using a mini Transblot Electrophoretic transfer cell system (Bio-Rad) filled with blotting buffer. Later, the membrane was blocked in a blocking solution (5% milk powder in 1X TBST (0.05% Tween in Tris Buffered Saline)) for 30 minutes at RT on a shaker.

### **3.1.3 Antibody staining of blots**

All antibodies were diluted in 5% milk powder in 1X TBST.

The membranes were incubated in the primary antibodies (Table 1) placed in falcon tubes on a rotator at RT for one hour or overnight at +4°C. Membranes were washed with TBST 3 times for 10 minutes to minimize the non-specific binding of antibodies, and subsequently incubated in horseradish peroxidase (HRP)-conjugated secondary antibodies (Table 2) for one hour at RT (on a shaker). After the incubation was over, the membranes were washed with TBST three times for 10 minutes. Chemiluminescence was detected using Clarity Western ECL Substrate (Bio-Rad, Cat. #170-5061), and images were acquired by Sapphire Biomolecular Imager (Azure Biosystems). For chemiluminescence detection, ECL kit solutions were mixed together with the ratio of 1:1, and small drops of the mixture were applied on a cling film. Then, the membranes were placed on the drops upside down for 3 to 5 minutes. Finally, the membranes were scanned by Sapphire, using different exposure times.

**Table 1. Primary antibodies used in Western blot.**

Primary antibodies	Produced in	Source	Cat No.	Dilution
Anti-Rab24 monoclonal antibody	Mouse	BD transduction	612174	1:1000
Anti-Rab24 polyclonal antibody	Rabbit	Proteintech	11445-1-AP	1:400
Anti-Tubulin E7 Monoclonal antibody	Mouse	DSHB	AB_1157911	1:5000

**Table 2. Secondary antibodies used in Western blot.**

Secondary antibodies	Produced in	Source	Antigen	Dilution
Anti-mouse HRP	Goat	Jackson Laboratory	Mouse IgG	1:20000 (WB)
Anti-rabbit HRP	Goat	Jackson Laboratory	Rabbit IgG	1:20000 (WB)

### 3.1.4 Western blot image analysis

To analyze Western blot images, each protein band's intensity was measured using ImageJ (Fiji) software. Initially, the grey mean value was checked through the "Analyze"> "Set the measurement" route in the software. Then the region of interest (ROI), which was the protein band at the first lane, was selected using the rectangle tool in Fiji through the following pathway: "Analyze"> "Gels"> "Select the first Lane pathway". The rectangle covered the protein band, for which we aimed to measure the intensity. For measuring other protein bands, we covered them by the same frame. The intensity was measured through running "Analyze"> "Gel"> "Plot Lanes". Once the plot appeared, the area under the curve was selected by wand (tracing) tools. The "result" window appeared, and intensity values were exported to an excel file. The normalization

was done by dividing antibody intensity by the loading control intensity. The loading control (tubulin) was stained and measured to normalize for difference in sample loading and transfer.

### ***3.2 Immunohistochemistry***

In immunohistochemical staining all the incubations took place in RT.

This method detects the antigen of interest in tissue sections employing immunological and chemical reactions. First, the sections were deparaffinized by submersion in xylene (2 x 10 minutes) and rehydrated by graded ethanol series from absolute to 50% ethanol (2 x 3 minutes) and water (2 x 5 minutes). Next, the sections were incubated in 3% hydrogen peroxide diluted in methanol for 20 minutes to block the endogenous peroxidase activity, then rinsed with water and 1X PBS (140 mM NaCl, 2.7 mM KCl, 4.6 mM Na<sub>2</sub>HPO<sub>4</sub> 2H<sub>2</sub>O, 1.5 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) for 5 minutes. For antigen retrieval, the sections were boiled in citrate buffer (10 mM Sodium citrate, 0.05% Tween 20, pH 6.0) for 15 minutes in a microwave oven (600W). Soon after boiling, the sections were cooled down to RT and washed in Tris-buffered saline (TBS, 150 mM NaCl, 50 mM Tris-Cl, pH 7.5). Permeabilization took place by placing the sections in TBST (TBS + 0.05% Tween 20) + 0.1% Triton X-100. Blocking was done for 1 hour using 5% normal goat serum for rabbit Rab24, diluted in TBS. Eventually, the sections were covered with the primary antibody (Table 3) diluted in TBS overnight at +4°C. The next day, sections were rinsed in PBS (3 x 5 minutes), incubated in secondary antibody (Table 4) at RT for 1 hour, and rinsed again with PBS (3 x 5 minutes). The sections were covered by enhancer solution (ABC Elite (Avidin/Biotin) Systems, Vector laboratories, #Cat No. PK-6101 diluted in PBS for 40 minutes, and subsequently, rinsed with PBS (3 x 5 minutes). The sections were covered with diaminobenzidine (DAB, Vector laboratories, #Cat No. SK-4100) diluted in PBS for 2.5 minutes, placed in MilliQ water, and finally counterstained with hematoxylin for 10 seconds. After rinsing with tap water, the sections were dehydrated by 70%, 95%, and 100% ethanol (2 x 3 minutes) and xylene (2 x 5 minutes). After dehydration, mounting was done using PERTEX mounting medium.

Image acquisition was done by Panoramic P1000 – 3DHISTECH Ltd slide scanner, at 40X magnification, with a resolution of 0.25 µm/pixel, for 15 multi-layers.

**Table 3. Primary antibodies used in immunohistochemistry.**

Primary antibodies	Produced in	Source	Cat No.	Dilution
Anti-Rab24 polyclonal antibody	Rabbit	Proteintech	11445-1-AP	1:400

**Table 4. Secondary antibodies used in immunohistochemistry.**

Secondary antibodies	Produced in	Source	Cat No.	Dilution
Biotinylated goat anti-rabbit IgG	Goat	Vector laboratories	PK-6101	1:200

### 3.2.1 Histological quantitative image analysis

#### Software installation

To set up a robust quantitative histological image analysis method, I combined the open-source QuPath with StarDist. To do this, TensorFlow, an open-source software math library for machine learning was required. It focuses mainly on training and inference of deep neural networks. Because TensorFlow is not included in the main QuPath distribution, I needed to build QuPath from the source (Building QuPath | Pete's blog).

Once the installation was complete, QuPath was used to follow the workflow described below:

#### 1. Opening the image

Images were opened by importing the image using the "File"> "open" path or by dragging and dropping the image onto QuPath's main window.

At the beginning of opening an image, the software asks for the image type. In this experiment, Hematoxylin and DAB were used as chromogens; therefore, H-DAB was selected as image type.

If there are more than one image to analyze, a project can be created, and all images can be opened in that project. For this, the user creates a new folder, clicks on the "project" on the left-hand pin, selects "create project" button and then chooses the newly created folder. The series of images can be added to the new project by following "Add images," "choose file," select the images, and "Import" pathway in the "Project" pin. One of the benefits of creating a project is that switching between related images is more straightforward. Besides, by clicking on the next image in the project, the software will ask if the user wants to save the changes done on the previous image. By clicking "yes," all the changes will be saved in a new folder created for the project into a data subfolder (Projects — QuPath 0.2.3 documentation).

## 2. Checking image properties

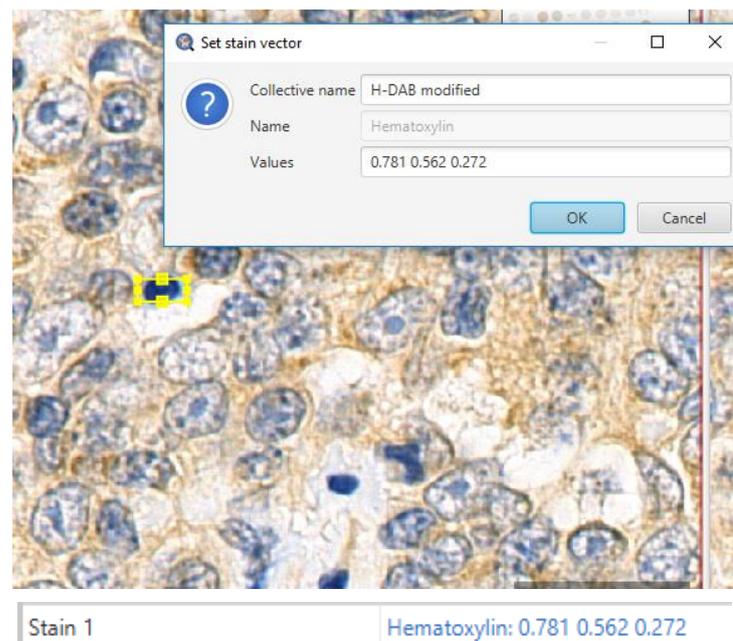
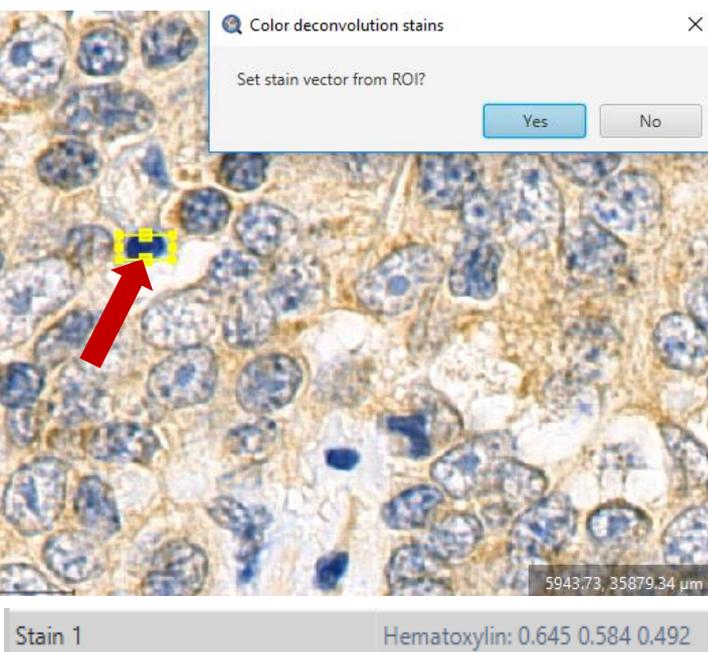
By clicking on the "image" tab on the left-hand pin, some of the primary metadata, including the image name, path, pixel size, image type, and stain vector, can be seen (Figure 5). By double-clicking on each of these, the user can adjust or change the parameters.

Project		Image		Annotations		Hierarchy		Workflow	
Name					Value				
Name					DigitalSlide_C1M_1S_1_20201210...				
URI					file:///utu.fi/taltio/Group-Eskelin...				
Pixel type					uint8 (rgb)				
Magnification					40.0				
Width					167424 px (20303.07 $\mu\text{m}$ )				
Height					391168 px (47457.88 $\mu\text{m}$ )				
Dimensions (CZT)					3 x 1 x 1				
Pixel width					0.1213 $\mu\text{m}$				
Pixel height					0.1213 $\mu\text{m}$				
Uncompressed size					183 GB				
Server type					OpenSlide				
Pyramid					1 2 4 8 16 32 64 128 256 512				
Metadata changed					No				
Image type					Brightfield (H-DAB)				
Stain 1					Hematoxylin: 0.659 0.585 0.473				
Stain 2					DAB: 0.262 0.483 0.835				
Stain 3					Residual: 0.494 -0.811 0.314				
Background					252 252 253				

**Figure 5. Image tab in QuPath.** In this tab, some of the basic metadata, such as pixel size, image size and magnification can be seen.

### 3. Setting stain vectors

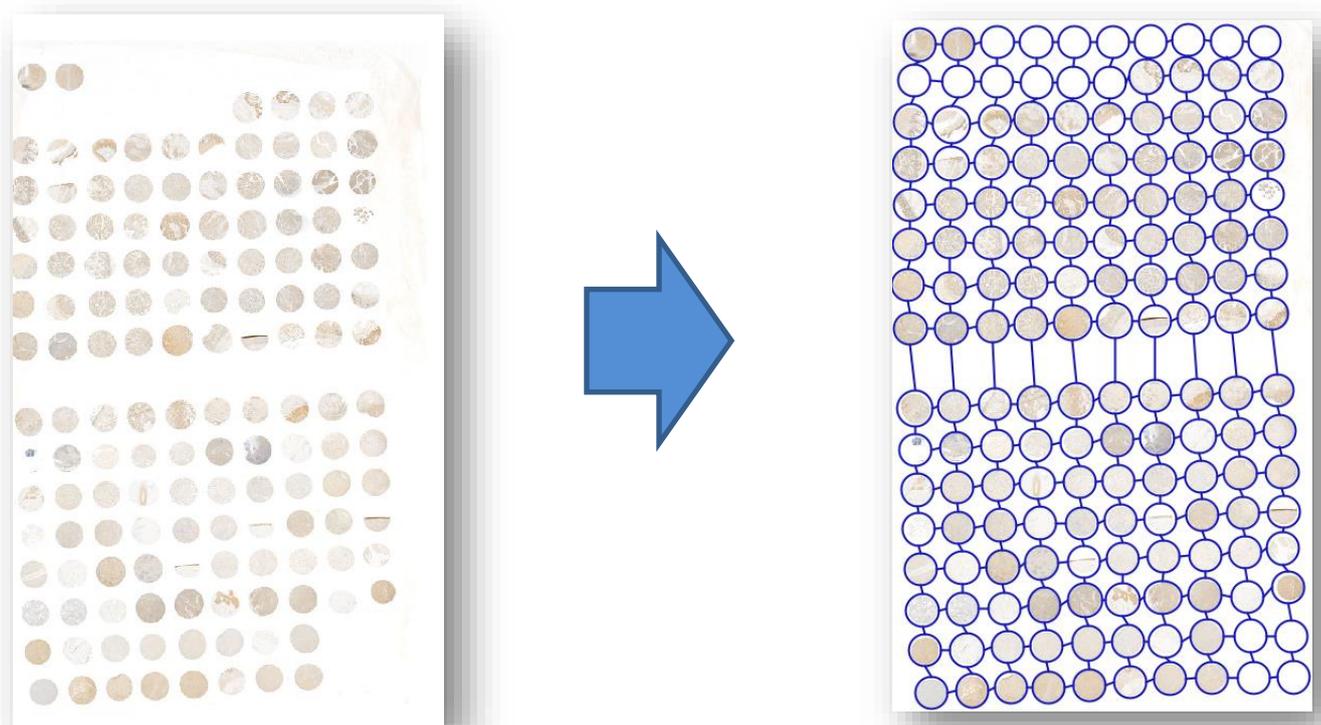
This step is not always necessary and does not affect the quantitative analysis significantly. However, stain vectors are the key to color deconvolution and are useful when the user wants to refine the stain estimate that QuPath uses by default. Stain vectors characterize which color values the blue and brown colors have for hematoxylin and DAB. To refine the values, a small rectangle was drawn that covered the nucleus area, which represented pure hematoxylin color. Then by double-clicking on "Stain" belonged to hematoxylin on the left-hand pin in the "image" tab and selecting "yes," slightly modified stain vector was seen (Figure 6). The same steps were repeated for DAB and background stain vector modification (Separating stains — QuPath 0.2.3 documentation).



**Figure 6. Image pin, setting stain vectors.** Image on the left demonstrates the stain estimation for hematoxylin that QuPath uses by default. Image on the right shows a slightly changed value for stain estimation after modification. The red arrow shows the annotation, which covered an area that represented hematoxylin color.

#### 4. Creating TMA grid

Selecting multiple tissue samples in TMAs at once is essential to receive data per tissue core in an arranged manner. Detecting cells in each tissue core individually is time-consuming. Therefore, an organized web of annotations (TMA grid) is required. To do this, first, a rectangle was drawn to select all tissue cores together. The script described in Text box 1 was then run to create annotations that cover tissue cores separately in one slide. The user can adjust the number of horizontal and vertical annotations depending on the number of tissue cores in each row and column in one slide. By adjusting size and position of the annotations manually, all cores were fully covered (Figure 7). All scripts used in the analysis are released with the author's permission and optimized for our project (Latest qupath topics - Image.sc Forum).



**Figure7. TMA selection.**

**Text box 1. TMA selection script (Groovy language)**

```

import qupath.lib.objects.PathObjects
import qupath.lib.objects.hierarchy.DefaultTMAGrid

// Enter the number of horizontal & vertical cores here
int numHorizontal = 10
int numVertical = 17
// Enter the core diameter, in millimetres
double diameterMM = 1.25

// Convert diameter to pixels
double diameterPixels = (diameterMM * 1000) /
getCurrentImageData().getServer().getPixelCalibration().getAveragedPixelSizeMicrons()

// Get the current ROI
def roi = getSelectedROI()

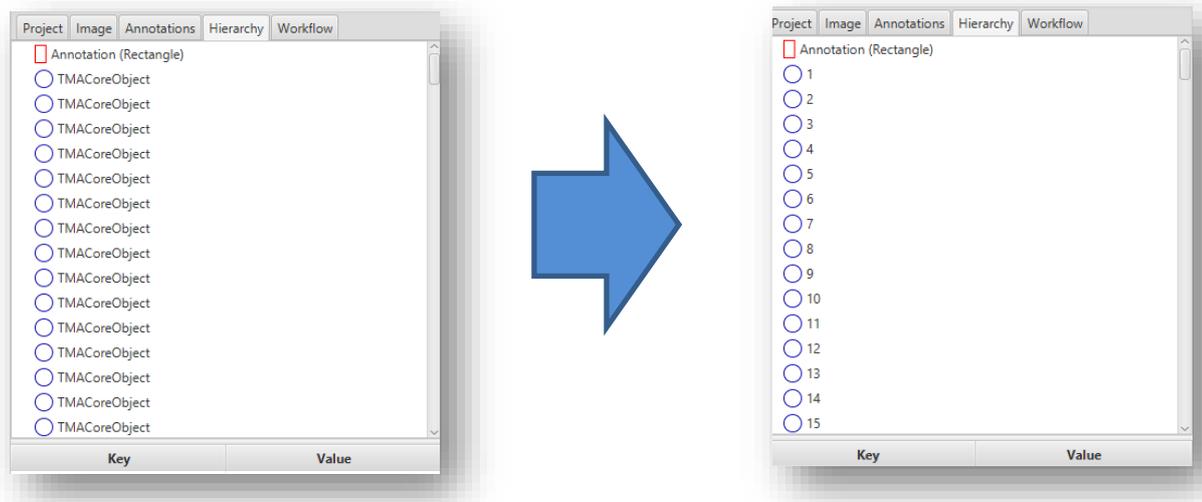
// Create the cores
def cores = []
double xSpacing = roi.getBoundsWidth() / numHorizontal
double ySpacing = roi.getBoundsHeight() / numVertical
for (int i = 0; i < numVertical; i++) {
    for (int j = 0; j < numHorizontal; j++) {
        double x = roi.getBoundsX() + xSpacing / 2 + xSpacing * j
        double y = roi.getBoundsY() + ySpacing / 2 + ySpacing * i
        cores << PathObjects.createTMACoreObject(x, y, diameterPixels, false)
    }
}

// Create & set the grid
def tmaGrid = new DefaultTMAGrid(cores, numHorizontal)
getCurrentHierarchy().setTMAGrid(tmaGrid)
relabelTMAGrid("1-6", "A-F", true)

```

## 5. Numbering of TMA cores

The tissue cores in the TMAs were numbered starting from 1, to easily distinguish the measurements from each other (Figure 8). The script described in Text box 2 was run to number the cores.



**Figure 8. TMAs before and after numbering the cores.**

Tissue core numbers can be seen in hierarchy pin.

### **Text box 2. Script for numbering each core in a TMA (Groovy language)**

```
getTMACoreList().eachWithIndex { c, i -> c.setName(i + 1 as String)}
fireHierarchyUpdate()
```

## 6. Commenting

Commenting on each tissue core was done by right-clicking the TMA grid and following the "TMA" > "Add TMA note" path. Comments can be seen through the Hierarchy tab.

## 7. Manual classification

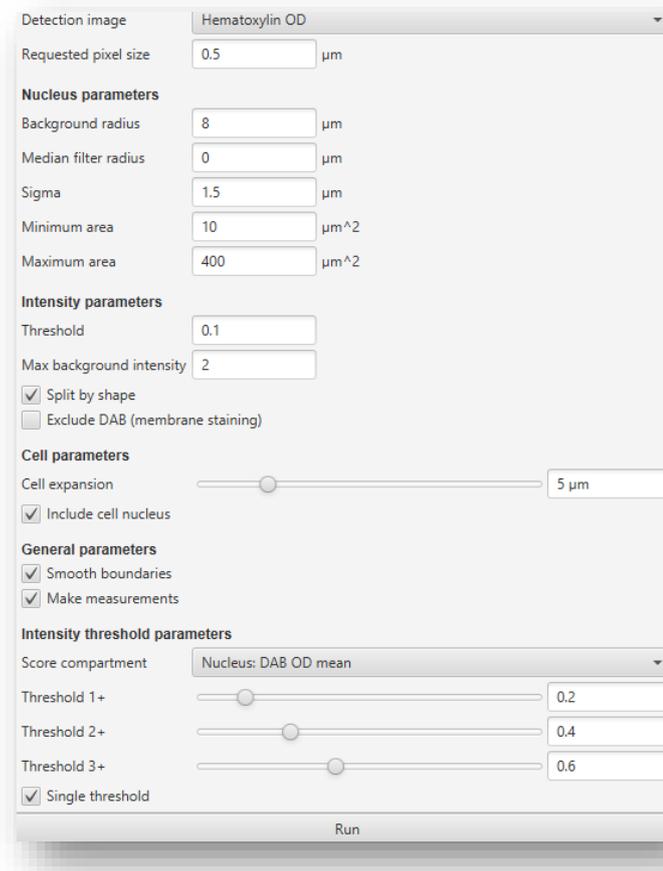
In QuPath, cells can be categorized into different classes such as tumor, stroma, immune cells, and positive/negative cells. Classifying a group of cells helps the user to analyze which cell types the positivity of staining is focused in. The classification was done using two classification strategies. First, manual classification was done after the commenting step, by annotating the region of interest (such as the tumor) manually. For this, QuPath's list of tools, including rectangle, brush, and wand tools, were used. Then the cells were classified through the following pathway: right click on the annotation > "Set class" > choose the category.

In this study, tissues were classified into four categories: epithelial tissue, connective tissue, connective tissue mixed with muscle, and muscle tissue. New classes can be defined by going to the left-hand pin, clicking on the Annotations button, and selecting "Add a class."

Train a cell classifier command was also applied to improve the manual classification by training the cells automatically. This command is explained in section 9. Note that shortcut Ctrl+L was used to access the list of commands and search bar (Cell classification — QuPath 0.2.3 documentation).

## 8. Positive cell detection

This step aimed to detect, count, and score the cells inside each tissue core and annotation. The aim was to score the staining intensity of Rab24 IHC reaction. To this end, the positive cell detection (Detecting positive cells | Pete's blog) command was run by searching in the list of commands (Ctrl+L) or by running the "Analyze" > "Cell analysis" > "Positive cell detection." The action brought up a dialog containing adjustable parameters related to cell detection. These parameters included detection image, pixel size, score compartment, and positivity classification at the bottom of the dialog box (Figure 9).



**Figure 9. Positive cell detection command in QuPath.**

Key parameters are described below:

### I. Choose detection image

The protein of interest, Rab24, was expected to be localized in the cytoplasm (brown DAB). "Hematoxylin OD" does not work well for analysis of the parts stained brown. It gives the brown color a low value after the stain separation, resulting in biased detection and over fragmentation of some nuclei. Instead, "optical density sum" was used, which uses all the brightness information regardless of the stain color, gives a higher value for positive nuclei, and thus improved the detection.

### II. Requested pixel size

This parameter defines the image resolution the QuPath will run the detection on. By clicking on the "image" button in the left-hand pin, image pixel size was shown. Our image pixel size was 0.1213  $\mu\text{m}$ , but we preferred not to run the detection on the actual resolution and used 0.5  $\mu\text{m}$  instead. Using this value showed a more accurate detection.

### III. Background radius and threshold

These parameters are relatively important when it comes to improving detection.

QuPath subtracts the background value from every pixel of the image during the detection. After that, it will use a threshold that defines the pixels that will pass that threshold and will be detected. For our measurement, the optimal value for background radius was 8. The threshold is explained more in section VI.

#### IV. Median filter radius and sigma

These parameters were used to fix the blurred or fragmented nuclei. For reducing the number of false fragmented nuclei, higher values need to be used. The optimal values for our analysis were 0.5 and 1.75 for median filter and sigma, respectively.

#### V. Minimum area and maximum area

These parameters define the smallest and largest nuclei that are allowed to be detected. We detected the nuclei between 5 and 1000  $\mu\text{m}^2$ .

#### VI. Threshold

Threshold works based on maximum background intensity. It tells QuPath not to apply the detection if the background is above a certain threshold. This is useful to avoid section folds. The optimal value for our purpose was 0.1.

#### VII. Split by shape box

This parameter works similar to the watershed, a transformation defined on binary images that separates adjacent nuclei from one another.

#### VIII. Exclude DAB box

By ticking this box, QuPath avoids detecting pixels where the DAB value is higher than the hematoxylin value. It is mainly used when the DAB staining is distributed outside the nucleus. This parameter was used in our experiment.

## IX. Cell expansion

Cell expansion estimates the cell size based upon distance from the nucleus and other nuclei's proximity. The optimized cell expansion value for this project was 6  $\mu\text{m}$ .

## X. Make a measurement box

Make a measurement was used in this study in order to get the measurements per cell.

## XI. Score compartment and thresholds

DAB OD mean as score compartment was selected. The optimal threshold values for scoring the cells were 0.06, 0.14, and 0.28 for 1+ (Weekly positive), 2+ (Moderately positive), and 3+ (Strongly positive), respectively. Below 0.06 was considered negative. These values were selected to train data after discussion with Professor Ilmo Leivo, a specialist in pathology.

Positive cell detection analysis can be applied to the tissues at once using the script in Text box 3.

### **Text Box 3. Positive cell detection script (Groovy language)**

The Positive cell detection script is shown as saved by the commands' history.

```
setImageType('BRIGHFIELD_H_DAB');

setColorDeconvolutionStains({'Name' : "H-DAB modified", "Stain 1" : "Hematoxylin", "Values 1" : "0.65892 0.58493 0.47294 ", "Stain 2" : "DAB", "Values 2" : "0.26211 0.4832 0.83535 ", "Background" : " 252 252 253 "});

runPlugin('qupath.imagej.detect.cells.PositiveCellDetection', '{"detectionImageBrightfield": "Optical density sum", "requestedPixelSizeMicrons": 0.5, "backgroundRadiusMicrons": 8.0, "medianRadiusMicrons": 0.5, "sigmaMicrons": 1.75, "minAreaMicrons": 5.0, "maxAreaMicrons": 1000.0, "threshold": 0.1, "maxBackground": 2.0, "watershedPostProcess": true, "excludeDAB": true, "cellExpansionMicrons": 6.0, "includeNuclei": true, "smoothBoundaries": true, "makeMeasurements": true, "thresholdCompartment": "Cytoplasm: DAB OD mean", "thresholdPositive1": 0.06, "thresholdPositive2": 0.14, "thresholdPositive3": 0.28, "singleThreshold": false}');
```

Even though we optimized positive cell detection command parameters to get the most reliable result, many false fragmented nuclei could not be excluded. To fix this, instead of using positive cell detection command, the Deep learning network StarDist was run in QuPath. To run StarDist on the image, the modified StarDist script described in Text box 4 was run.

#### **Text box 4. Script for running StarDist (Groovy language)**

```
import qupath.tensorflow.stardist.StarDist2D

// Specify the model directory (you will need to change this!)

def pathModel = "C:\\Users\\Safaze\\Desktop\\he_heavy_augment"

def stardist = StarDist2D.builder(pathModel)

    .threshold(0.1)      // Probability (detection) threshold
//   .channels('DAPI')   // Select detection channel

    .normalizePercentiles(1, 99) // Percentile normalization

    .pixelSize(0.5)     // Resolution for detection

    .cellExpansion(5.0) // Approximate cells based upon nucleus expansion

    .cellConstrainScale(1.5) // Constrain cell expansion using nucleus size

    .measureShape()     // Add shape measurements

    .measureIntensity() // Add cell measurements (in all compartments)

    .includeProbability(true) // Add probability as a measurement (enables later filtering)

    .build()

// Run detection for the selected objects

def imageData = getCurrentImageData()

def pathObjects = getSelectedObjects()

selectTMACores()

if (pathObjects.isEmpty()) {

    Dialogs.showMessageDialog("StarDist", "Please select a parent object!")

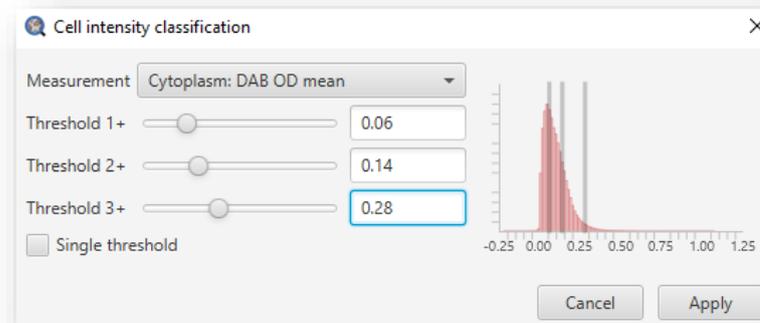
    return

}

stardist.detectObjects(imageData, pathObjects)

println 'Done!'
```

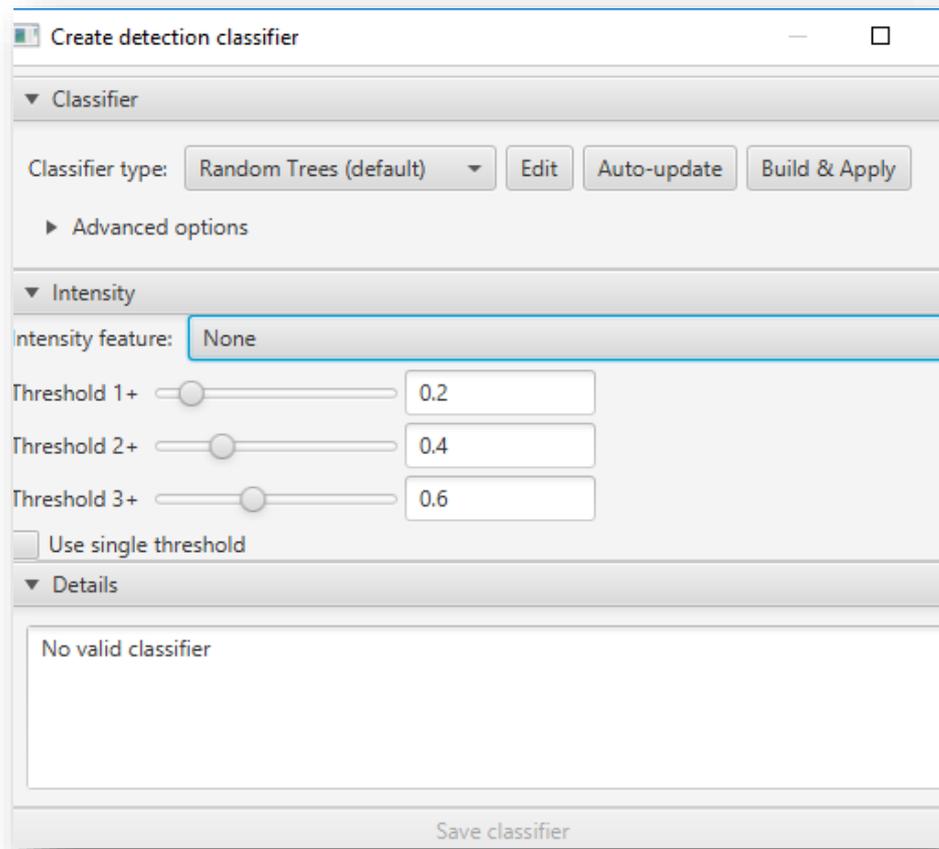
StarDist does not grade the cells based on their intensity. To score the cells, cell intensity classification (Figure 10) was run through the following pathway: "Classify" > "Object classification" > "Set cell intensity classification." The optimal values for scoring the cells were similar to those used in positive cell detection.



**Figure 10. Intensity classification.** Post-intensity classification main window in StarDist.

## 9. Training a cell classifier

Trainable Classification was done at this stage to improve the manual Classification. Classifier dialog was brought up in QuPath, running "Classify" > "Object classification" > "Older classifiers" > "Create detection classifier" (Figure 11). Next, the annotation tools were applied to give QuPath examples of different cell types in each class. For this, 5 to 7 cells were annotated and classified into different classes similar to section 7. Then "Auto-update" button in the classifier dialog was selected in order to classify all the cells across the image automatically (Cell classification — QuPath 0.2.3 documentation).



**Figure 11. Cell classifier.** Automated cell classification is done, using create detection classifier.

## 10. Measurement tables and exporting the data

Ultimately, the output was seen in TMA/annotation measurement tables through the following path: "Measure"> "Show annotations/TMA measurements."

The results were exported using the save button in "TMA/annotation measurements" tables. Annotation measurements were sorted into different classes that were defined previously. The script in Text box 5 was run before exporting the result in order to connect each annotation to the corresponding tissue core (Exporting measurements — QuPath 0.2.3 documentation).

**Text box 5. Script for assigning an annotation to be below a TMA core based on its centroid location (Groovy language)**

```
def hierarchy = getCurrentHierarchy()
def tmaGrid = getCurrentHierarchy().getTMAGrid()
def unassigned = getAnnotationObjects().findAll {it.getParent().isRootObject()}
removeObjects(unassigned, true)

for (annotation in unassigned) {
    def roi = annotation.getROI()
    def core = PathObjectTools.getTMACoreForPixel(tmaGrid,roi.getCentroidX(),
roi.getCentroidY())
    hierarchy.addPathObjectBelowParent(core, annotation, false)
}

fireHierarchyUpdate()
print 'Done!'
```

## 4. Results

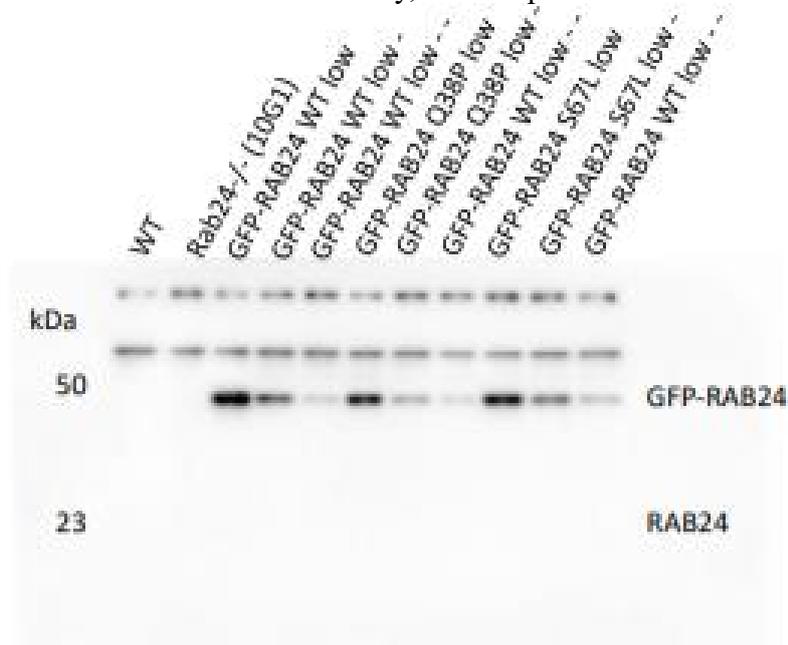
### ***4.1 Rab24 antibody validation***

Rabbit anti-Rab24 (polyclonal, Proteintech, Cat No. #11445-1-AP) was used for the IHC since it had been used for IHC in a published paper. Proteintech website shows a Western blot of mouse brain lysate, stained with this antibody, where only one band is visible at the expected position of 23 kDa (RAB24 Antibody 11445-1-AP | Proteintech). The website also shows images of IHC staining with this antibody. Western blotting of wildtype and Rab24 knockout Neuro-2a cells was used to validate the specificity of the antibody.

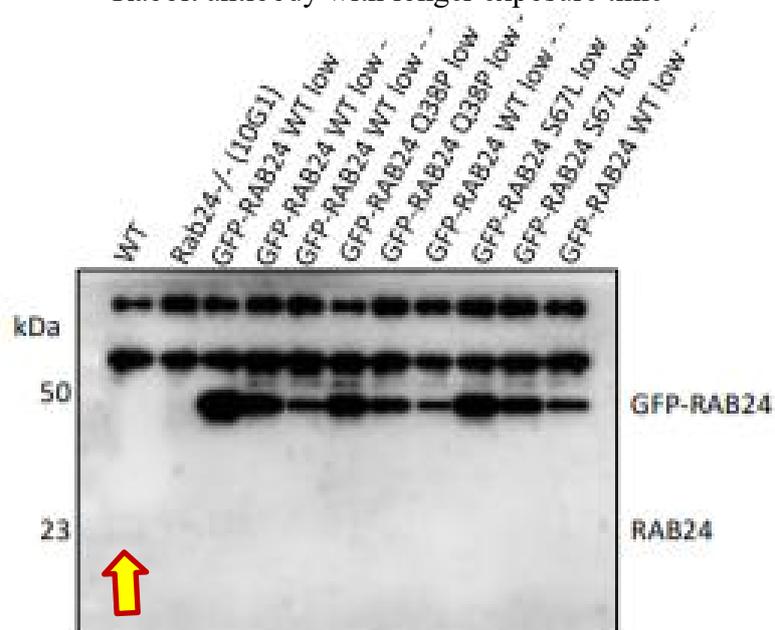
#### **4.1.1 Western blot for antibody validation**

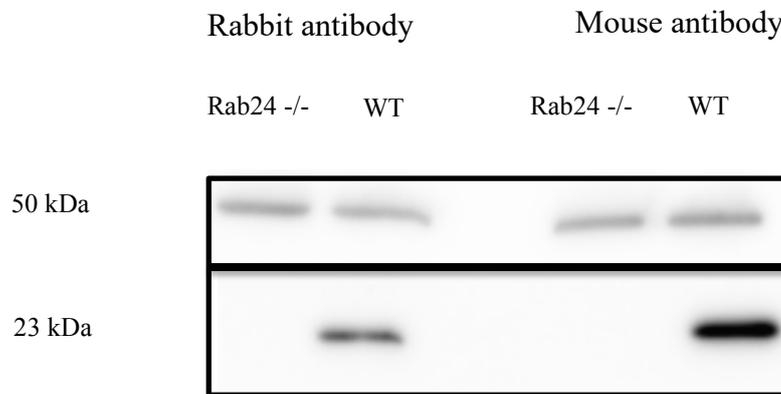
WT and Rab24 knockout (Rab24<sup>-/-</sup> clone10G1), and GFP-tagged Rab24 overexpressing Neuro-2a cell lines were used as samples. Unexpectedly, in addition to the expected Rab24 band at 23 kDa in WT cells, we saw two additional bands above 50 kDa in both WT and Rab24 knockout cells (Figure 12A, B). The nature of these unspecific bands is currently unknown. Since the rabbit anti-Rab24 gave unspecific bands in Western blotting, we tested a mouse anti-Rab24 (monoclonal, BD transduction, Cat No. #612174), in addition to the rabbit anti-Rab24, in Western blot with the Neuro-2a cell lines. In this case, the upper part of the blot was used for staining with other antibodies, so the unspecific bands over 50 kDa are not visible. Nevertheless, both rabbit anti-Rab24 and mouse anti-Rab24 show the endogenous Rab24 band at 23 kDa in wildtype cells, and no band in the Rab24<sup>-/-</sup> cells (Figure 12B). The mouse anti-RAB24 used for the immunoblotting is not recommended for IHC by the manufacturer. Our preliminary experiments confirmed that it is not suitable for IHC of paraffin embedded sections.

Rabbit antibody, short exposure



Rabbit antibody with longer exposure time



**B)**

**Figure12) Western blot for antibody validation.** Rabbit anti-Rab24 (panels A, B) and mouse anti-Rab24 (panel B) were used to staining Western blots of wildtype (WT), Rab24<sup>-/-</sup> (10G1), and GFP-Rab24 overexpressing Neuro-2a cell extracts. The minus signs indicate medium (-) and low (--) expression levels of GFP-Rab24. Note that the GFP-Rab24 expressing cells do not have endogenous Rab24. Yellow arrow in panel A shows the weak band for endogenous Rab24 in WT cells at 23 kDa.

## ***4.2 TMA quantification***

After performing IHC and image acquisition on the paraffin sections of the TMAs, StarDist was run on QuPath to analyze the images. The output produced data regarding the number of detected cells in TMA cores and annotations, categorized in different staining intensity groups. The measurement tables also provided an index called "histo" score or H-score. H-score is a quantitative evaluation index that takes different staining intensities viewed at various magnifications into account. This is done by adding the intensity to the percentage of positive cells. H-score in QuPath was calculated for all the detected cells in each cancer core and also for each tissue type in each cancer core classified in different subgroups known as epithelial, connective, connective and muscle, and muscle tissues.

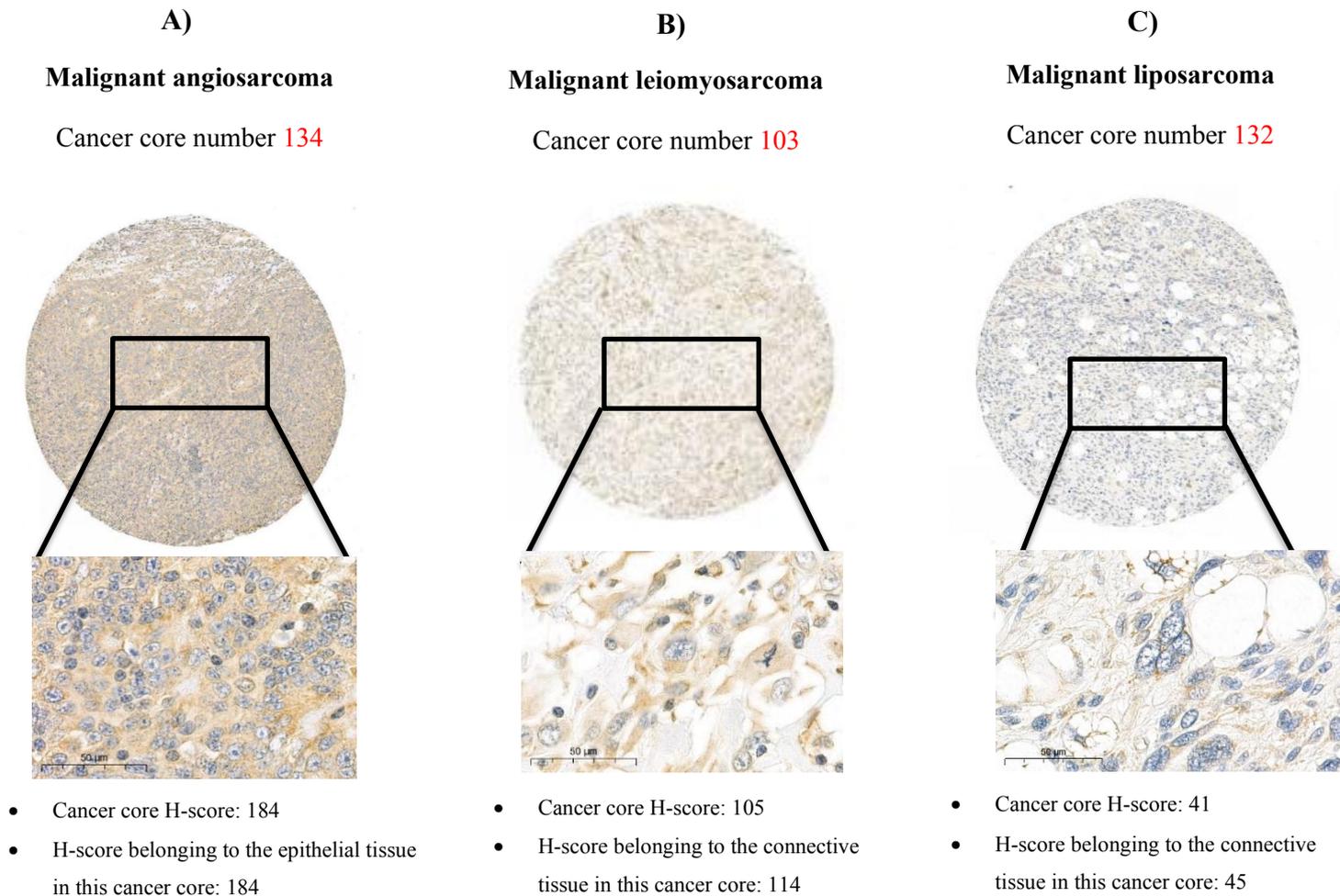
H-score is calculated using the following formula:

$$\text{H-score} = [1 \times (\% \text{ cells } 1+) + 2 \times (\% \text{ cells } 2+) + 3 \times (\% \text{ cells } 3+)].$$

The H-score method assigns a score of 0 - 300 for each tissue core and annotation. The discriminatory threshold can be set at a certain value based on the staining and the aim of the study. It should be noted that in addition to overexpression of Rab24, very low Rab24 expression could also be potentially valuable as a biomarker. Therefore, much importance was placed on H-scores calculated from all cells in each cancer core, significantly higher or lower than our on-slide liver control (H-score 164). For this reason, cancer tissues with H-scores more than 180 and less than 90 were selected as potential candidates for further investigation. Cancer tissues with H-scores < 40 and > 210 were excluded due to the biologically plausible cutoffs (Mazières et al., 2013; April 10, 2015 - Calculating H-Score - The ASCO Post; QuPath H-Scoring - Image Analysis - Image.sc Forum). Most of the tissue cores belonging to the angiosarcoma, stomach GIST, and rectal NEC were ranked the highest, and liposarcoma and fibrosarcoma were ranked the lowest based on the H-scores. Furthermore, there was less fluctuation in the H-scores among different patients in these 5 cancer types.

#### 4.2.1 Classification of tissues based on staining intensity.

Examples of intensely stained, moderately stained, and weakly stained TMAs are shown in Figure 13.



**Figure 13. TMA cores staining intensity comparison.** Three tissue cores, categorized as strongly stained (A), moderately stained (B), and weakly stained (C), are shown. Tissue core belonging to malignant angiosarcoma was stained intensely using anti-Rab24, while liposarcoma was stained weakly. H scores are in agreement with the morphology of cells in the tissue cores. One cancer core usually contains more than one tissue type; for instance, there may be both connective tissue and epithelial tissue in one cancer core. Cancer core H-score is calculated from all scored cells in each cancer core. Tissue H-score belongs only to the cells in one tissue type in a cancer core. Cancer core H-scores were used to compare cancer cores with each other in terms of Rab24 expression quantitatively. Tissue H-scores were used to see which tissue type (connective, epithelial, or muscle) was stained more intensely. Tissues in one cancer core can have similar or different H-scores from their corresponding cancer core depending on the tissue heterogeneity in that core. Scale bar: 50 µm.

**Table 5. Percentage of all scored cells in the cancer cores above categorizing in different intensity groups.**

<b>Percentage of scored cells</b>	<b>Malignant angiosarcoma</b>	<b>Malignant leiomyosarcoma</b>	<b>Malignant liposarcoma</b>
<b>1+</b>	<b>15.74</b>	<b>64.55</b>	<b>39.71</b>
<b>2+</b>	<b>82.67</b>	<b>18.94</b>	<b>0.60</b>
<b>3+</b>	<b>1.08</b>	<b>0.76</b>	<b>0.05</b>
<b>negative</b>	<b>0.46</b>	<b>15.71</b>	<b>59.64</b>

H-scores obtained for each cancer core and most of the tissue types corresponding to those cancer cores in the analysis are presented in Tables 6 and 7.

**Table 6. QuPath analysis.** H-scores was calculated using all scored cells in each cancer core. Each row corresponds to one cancer core.

Cancer core numbers in the analysis	Number of patients	Note	H-score
1		Liver control 1	164
2		Liver control 2	162
17	3	Low-grade appendic mucinous neoplasms	128
18		Low-grade appendic mucinous neoplasms	152
19		Low-grade appendic mucinous neoplasms	84
20		Low-grade appendic mucinous neoplasms	108
21		Low-grade appendic mucinous neoplasms	140
22		Low-grade appendic mucinous neoplasms	150
23	3	Benign esophagus epidermoid carcinoma	180
24		Esophagus epidermoid carcinoma	141
25		Esophagus epidermoid carcinoma	120
26		Benign esophagus epidermoid carcinoma	152
27		Esophagus epidermoid carcinoma	113
28		Esophagus epidermoid carcinoma	101
29		Benign esophagus epidermoid carcinoma	51
30		Esophagus epidermoid carcinoma	183
31		Esophagus epidermoid carcinoma	175
32	3	Esophagus adenocarcinoma	147
33		Esophagus adenocarcinoma	163
34		Esophagus adenocarcinoma	119
35		Esophagus adenocarcinoma	126
36		Esophagus adenocarcinoma	89
37		Esophagus adenocarcinoma	157
38	1	Benign rectal NEC	128
39		Rectal NEC	206
40		Rectal NEC	195
41	3	Benign rectal adenocarcinoma	165
42		Rectal adenocarcinoma	132
43		Rectal adenocarcinoma	158
44		Rectal adenocarcinoma	104
45		Rectal adenocarcinoma	218
46		Rectal adenocarcinoma	136
47		Rectal adenocarcinoma	110
48	1	Small intestine NEC	151
49		Small intestine NEC	146
50	3	Benign small intestine Neuroendocrine Tumor (NET) grade 2	172
51		Small intestine NET grade 2	121
52		Small intestine NET grade 2	163

("Table 6. (Continued).")

Cancer core number	Number of patients	Note	H-score
53	...	Benign small intestine NET grade 2	130
54		Small intestine NET grade 2	147
55		Small intestine NET grade 2	124
56		Benign small intestine NET grade 2	137
57		Small intestine NET grade 2	124
58		Small intestine NET grade 2	128
59	3	Benign small intestine NET grade 1	175
60		Small intestine NET grade 1	156
61		Small intestine NET grade 1	177
62		Benign small intestine NET grade 1	149
63		Small intestine NET 1	145
64		Small intestine NET grade 1	134
65		Benign small intestine NET grade 1	68
66		Small intestine NET grade 1	140
67		Small intestine NET grade 1	140
68	1	Ileum neuroendocrine carcinoma	149
69		Ileum neuroendocrine carcinoma	138
70	2	Benign stomach and small intestine NEC grade 3	144
71		Stomach and small intestine NEC 3	182
72		Stomach and small intestine NEC 3	139
73		Stomach and small intestine NEC 3	146
74		Stomach and small intestine NEC 3	151
75	3	Benign stomach diffuse adenocarcinoma	201
76		Stomach diffuse adenocarcinoma	122
77		Stomach diffuse adenocarcinoma	116
78		Stomach diffuse adenocarcinoma	107
79		Stomach diffuse adenocarcinoma	133
80		Benign stomach diffuse adenocarcinoma	173
81		Stomach diffuse adenocarcinoma	135
82		Stomach diffuse adenocarcinoma	94
83	3	Stomach intestinal adenocarcinoma	109
84		Stomach intestinal adenocarcinoma	163
85		Benign stomach intestinal adenocarcinoma	181
86		Stomach intestinal adenocarcinoma	101
87		Stomach intestinal adenocarcinoma	108
88		Benign stomach intestinal adenocarcinoma	134
89		Stomach intestinal adenocarcinoma	123
90		Stomach intestinal adenocarcinoma	138

("Table 6. (Continued).")

Cancer core number	Number of patients	Note	H-score
91	3	Rhabdomyosarcoma	54
92		Rhabdomyosarcoma	96
93		Benign Rhabdomyosarcoma	60
94		Rhabdomyosarcoma	80
95		Rhabdomyosarcoma	78
96		Rhabdomyosarcoma	148
97		Rhabdomyosarcoma	107
98	4	Benign leiomyosarcoma	55
99		Leiomyosarcoma	114
100		Leiomyosarcoma	107
101		Benign leiomyosarcoma	127
102		Leiomyosarcoma	115
103		Leiomyosarcoma	105
104		Benign leiomyosarcoma	83
105		Leiomyosarcoma	74
106		Leiomyosarcoma	66
107		Leiomyosarcoma	82
108		Leiomyosarcoma	59
109	3	Malignant fibrose histocytoma	131
110		Malignant fibrose histocytoma	122
111		Malignant fibrose histocytoma	46
112		Malignant fibrose histocytoma	124
113		Malignant fibrose histocytoma	127
114		Benign malignant fibrose histocytoma	33
115		Malignant fibrose histocytoma	75
116	Malignant fibrose histocytoma	57	
117	3	Benign synovial sarcoma	66
118		Synovial sarcoma	163
119		Synovial sarcoma	118
120		Synovial sarcoma	156
121		Synovial sarcoma	104
122		Benifn synovial sarcoma	41
123		Synovial sarcoma	194
124		Synovial sarcoma	108
125	3	Benign liposarcoma	96
126		Liposarcoma	72
127		Liposarcoma	87
128		Liposarcoma	67

("Table 6. (Continued).")

Cancer core number – Tissue class	The number of patients	Note	H-score
129		Liposarcoma	71
130		Benign liposarcoma	69
131		Liposarcoma	60
132		Liposarcoma	41
133	3	Benign Angiosarcoma	49
134		Angiosarcoma	184
135		Angiosarcoma	173
136		Benign Angiosarcoma	186
137		Angiosarcoma	173
138		Angiosarcoma	155
139		Benign Angiosarcoma	34
140		Angiosarcoma	222
141		Angiosarcoma	197
142		3	Benign Fibrosarcoma
143	Fibrosarcoma		50
144	Fibrosarcoma		78
145	Fibrosarcoma		102
146	Fibrosarcoma		57
147	Benign Fibrosarcoma		34
148	Fibrosarcoma		101
151	Fibrosarcoma		65
152	3	stomach GIST	206
153		Benign stomach GIST	165
154		stomach GIST	181
155		stomach GIST	203
156		Benign stomach GIST	92
157		stomach GIST	119
158		stomach GIST	124

**Table 7. H-scores in different tissues.** The table shows H-scores belonging to epithelial, connective and muscle tissue in each of the tissue cores.

Cancer core number - Tissue class	H-score	TMA number - class	H-score	TMA number - class	H-score
140 - Connective	227	145 - Connective	108	147 - Connective	42
152 - Connective	209	148 - Connective	105	117 - Connective	35
155 - Connective	209	127 - Connective	92	103 - Connective	114
123 - Connective	202	36 - Connective	89	114 - Connective	32
141 - Connective	202	156 - Connective	87		
154 - Connective	186	144 - Connective	86		
33 - Connective	175	136 - Connective	85		
153 - Connective	171	126 - Connective	81		
118 - Connective	166	115 - Connective	81		
109 - Connective	138	128 - Connective	76		
35 - Connective	137	130 - Connective	76		
42 - Connective	137	65 - Connective	75		
32 - Connective	134	129 - Connective	74		
113 - Connective	132	111 - Connective	70		
112 - Connective	132	17 - Connective	66		
34 - Connective	131	116 - Connective	66		
158 - Connective	129	151 - Connective	65		
110 - Connective	128	146 - Connective	65		
119 - Connective	125	18 - Connective	65		
157 - Connective	123	125 - Connective	63		
45 - Connective	122	131 - Connective	59		
120 - Connective	122	143 - Connective	55		
21 - Connective	120	133 - Connective	54		
124 - Connective	116	142 - Connective	46		
23 - Connective	111	132 - Connective	45		
121 - Connective	110	122 - Connective	43		

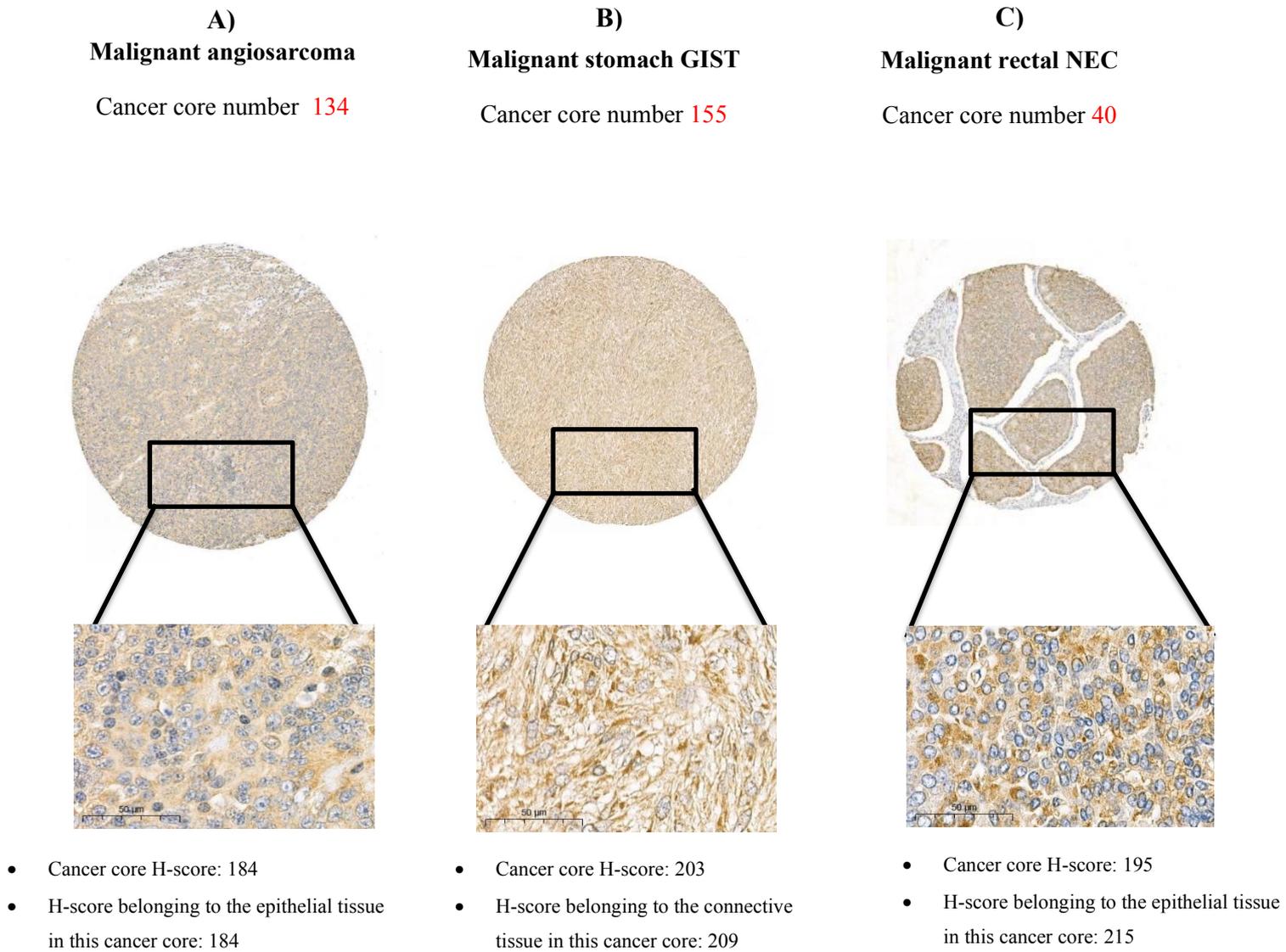
("Table 7. (Continued).")

<b>Cancer core number – Tissue class</b>	<b>H-score</b>	<b>TMA number - class</b>	<b>H-score</b>
96 - Connective and muscle	151	18 - Epithelial	259
81 - Connective and muscle	141	22 - Epithelial	253
89 - Connective and muscle	133	17 - Epithelial	253
76 - Connective and muscle	127	136 - Epithelial	253
97 - Connective and muscle	123	80 - Epithelial	247
102 - Connective and muscle	118	45 - Epithelial	241
79 - Connective and muscle	117	23 - Epithelial	228
100 - Connective and muscle	114	39 - Epithelial	221
83 - Connective and muscle	111	40 - Epithelial	215
90 - Connective and muscle	110	88 - Epithelial	209
22 - Connective and muscle	106	75 - Epithelial	206
20 - Connective and muscle	105	50 - Epithelial	183
78 - Connective and muscle	103	134 - Epithelial	184
80 - Connective and muscle	101	71 - Epithelial	181
92 - Connective and muscle	99	79 - Epithelial	178
82 - Connective and muscle	97	137 - Epithelial	178
19 - Connective and muscle	95	135 - Epithelial	173
77 - Connective and muscle	93	21 - Epithelial	171
101 - Connective and muscle	93	37 - Epithelial	156
94 - Connective and muscle	90	138 - Epithelial	150
107 - Connective and muscle	86	72 - Epithelial	139
95 - Connective and muscle	85		
104 - Connective and muscle	76		
105 - Connective and muscle	76		
108 - Connective and muscle	72		
106 - Connective and muscle	71		
88 - Connective and muscle	71		
98 - Connective and muscle	53		

("Table 7. (Continued).")

<b>Cancer core number – Tissue class</b>	<b>H-score</b>
78 - Muscle	139
21 - Muscle	137
99 - Muscle	120
51 - Muscle	109
86 - Muscle	98
87 - Muscle	88
93 - Muscle	66

Lastly, five cancer types were selected as candidates for further investigation in the future. Angiosarcoma, stomach GIST and rectal NEC were selected due to their high H-score (Figure 14). Liposarcoma and fibrosarcoma were also selected due to their low H-scores (Figure 15).



**Figure 14. IHC staining of selected cancer cores with high H-scores.** Most cancer cores in three cancer types, angiosarcoma (A), stomach GIST (B), and rectal NEC (C), were ranked higher than 180 by the H-score index. Further information can be found in the figure legend 13. Scale bar: 50  $\mu\text{m}$ .

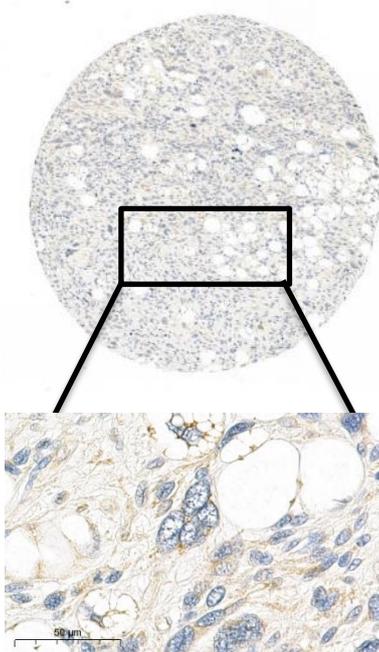
**Table 8. Percentage of all scored cells in cancer cores above categorizing in different intensity groups.**

Percentage of scored cells	Malignant angiosarcoma	Malignant stomach GIST	Malignant rectal NEC
1+	15.74	6.32	13.89
2+	82.67	83.72	53.14
3+	1.08	9.81	24.91
negative	0.46	0.15	8.07

A)

**Liposarcoma**

Cancer core number 132



- Cancer core H-score: 41
- H-score belonging to the connective tissue in this cancer core: 45

B)

**Fibrosarcoma**

Cancer core number 143



- Cancer core H-score: 50
- H-score belonging to the connective tissue in this cancer core: 55

**Figure 15. Selected cancer cores with low H-scores.** Most cancer cores in two cancer types, liposarcoma (A) and fibrosarcoma (B), were ranked lower than 90 by the H-score index. Only little fluctuation was observed among different patient's H-score in each of these two cancer types. Further information can be found in the figure legend 13. Scale bar: 50  $\mu$ m.

**Table 9. Percentage of all scored cells in cancer cores above categorizing in different intensity groups.**

Percentage of scored cells	Malignant liposarcoma	Malignant fibrosarcoma
1+	39.71	46.23
2+	0.60	1.59
3+	0.05	0.04
negative	59.64	52.14

## 5. Discussion

In this project, we screened Rab24 expression in TMA samples containing cores from 21 different human cancers. Collaboration with Professor Ilmo Leivo, an expert in pathology, and implementation of state-of-the-art quantification techniques simultaneously made this study a building block for future discoveries. An instant nuclei segmentation method is urgently needed for accurate quantification of histological staining. Therefore, I developed image analysis pipelines that provided a reliable and precise quantification.

Initially, I detected and scored the cells using "positive cell detection" command in QuPath. This command was not efficient for analysis even though I optimized the parameters, as it was developed mainly to detect the positive staining in the nuclei, while Rab24 is localized mainly in the cytoplasm. The analysis using "positive cell detection" command resulted in many false detections and over segmentations. Hence, I combined the QuPath with a deep learning network known as StarDist that is well-adapted to the typical shape of the nuclei to detect and score the cells. This method showed very promising results and minimized the errors to a great degree.

The following steps were taken in the project:

- Checking antibody validity
- Performing IHC on the tissue sections
- Discussion with the pathologist
- Image analysis and quantification

Using the rabbit anti-Rab24, we detected unknown positive bands in Western blots of the Neuro-2a cell lines, including Rab24  $-/-$  cells. This suggests that the antibody may give unspecific signals in IHC staining as well. Therefore, the results reported for the different cancer types should be confirmed using another validated Rab24 antibody.

It is important to point out that IHC using DAB as chromogen has some limitations as a quantitative method. This is because antigen/antibody amplification steps used during the IHC staining function in a non-linear fashion, and because DAB is not a true absorber of light. To ensure that results are sufficiently specific and sensitive, I implemented careful calibration using on-slide control tissues that showed low protein expression levels (Adding quality to your qualitative IHC | Medical Laboratory Observer). The analysis revealed angiosarcoma, stomach GIST, rectal NEC, liposarcoma, and fibrosarcoma as interesting cancer types for future studies.

As identification of cancer cells and benign cells in histology samples is challenging, further discussions with pathology experts are needed regarding the selected cancer types in order to draw firm conclusions on Rab24 expression in cancer and normal cells.

The findings also need to be investigated in a higher number of samples from a higher number of patients to gather data about the possible prognostic value of Rab24 expression in these cancers. Analyzing additional samples will also reveal whether there is a correlation between patients' survival and RAB24 expression.

The path toward success:

Cancer cores often contained multiple tissue types; therefore, I used manual classification and cell training classifier in QuPath to determine the tissue types and their H-scores individually. Classification revealed which tissue types were stained more positive compared to others. Despite the advanced "cell training classifier" command in open-source products such as QuPath and ImageJ-Fiji software, distinguishing different cell types, such as epithelial, connective, and muscle cells, is quite a challenge. "Training the cell classifier" command showed a high accuracy to classify cells that differ in intensity, but it cannot recognize the differences in the morphology yet. Thus, manual classification was better suited for this task. This issue can be developed further in the future.

## 6. Acknowledgement

I would like to express my gratitude to my supportive supervisor Professor Eeva-Liisa Eskelinen for giving me the opportunity to conduct my master's thesis project in her group.

Special thanks to Mauricio Ramm, my second supervisor, without whom this project's completion would not have been possible. I appreciate his patience and guidance and thank him for giving me motivation and confidence to complete the tasks.

I would like to extend my gratitude to Professor Ilmo Leivo for providing us with the pathological information required for our project.

I would also like to thank the coordinators, Raili Kronström and Joanna Pylvänäinen, for being there and solving many problems.

Big thanks to Pete Bankhead, Elnaz Fazeli, and Laura Murphy for helping me deal with the image analysis practical matters.

I would like to express my appreciation to Professor Pekka Hänninen, Dean of the Medical faculty, for his outstanding management skills and for providing me with a safe environment to conduct my research in the current corona pandemic.

Last but not least, I would like to thank my family, especially my parents and my dear sister, for believing in me and supporting me to continue my studies in Finland.

## 7. References

Adding quality to your qualitative IHC | Medical Laboratory Observer. URL: <https://www.mlo-online.com/diagnostics/biomarkers/article/13008492/adding-quality-to-your-qualitative-ihc>

Amaya, C., R.D. Militello, S.D. Calligaris, and M.I. Colombo. 2016. Rab24 interacts with the Rab7/Rab interacting lysosomal protein complex to regulate endosomal degradation: AMAYA et al. *Traffic*. 17:1181–1196. doi:10.1111/tra.12431.

Ao, X., L. Zou, and Y. Wu. 2014. Regulation of autophagy by the Rab GTPase network. *Cell Death Differ*. 21:348–358. doi:10.1038/cdd.2013.187.

April 10, 2015 - Calculating H-Score - The ASCO Post. URL: <https://ascopost.com/issues/april-10-2015/calculating-h-score/>

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, and P.W. Hamilton. 2017. QuPath: Open source software for digital pathology image analysis. *Sci. Rep.* 7:16878. doi:10.1038/s41598-017-17204-5.

Building QuPath | Pete's blog. URL: <https://petebankhead.github.io/qupath/2019/10/01/building-qupath.html>

Cell classification — QuPath 0.2.3 documentation. URL: [https://qupath.readthedocs.io/en/latest/docs/tutorials/cell\\_classification.html](https://qupath.readthedocs.io/en/latest/docs/tutorials/cell_classification.html)

Chen, Z., X. Wang, R. Liu, L. Chen, J. Yi, B. Qi, Z. Shuang, M. Liu, X. Li, S. Li, and H. Tang. 2017. KDM4B-mediated epigenetic silencing of miRNA-615-5p augments RAB24 to facilitate malignancy of hepatoma cells. *Oncotarget*. 8:17712–17725. doi:10.18632/oncotarget.10832.

COSMIC | Catalogue of Somatic Mutations in Cancer. URL: <https://cancer.sanger.ac.uk/cosmic>

Detecting positive cells | Pete's blog. URL: <https://petebankhead.github.io/qupath/tips/2018/03/22/setting-positive.html>

Ding, J., G. Soule, J.H. Overmeyer, and W.A. Maltese. 2003. Tyrosine phosphorylation of the Rab24 GTPase in cultured mammalian cells. *Biochem. Biophys. Res. Commun.* 312:670–675. doi:10.1016/j.bbrc.2003.10.171.

Dong, S., C. Aguirre-Hernandez, A. Scrivo, C. Eliscovich, E. Arias, J.J. Bravo-Cordero, and A.M. Cuervo. 2020. Monitoring spatiotemporal changes in chaperone-mediated autophagy in vivo. *Nat. Commun.* 11:645. doi:10.1038/s41467-019-14164-4.

Exporting measurements — QuPath 0.2.3 documentation. URL: [https://qupath.readthedocs.io/en/latest/docs/tutorials/exporting\\_measurements.html](https://qupath.readthedocs.io/en/latest/docs/tutorials/exporting_measurements.html)

Expression of RAB24 in cancer - Summary - The Human Protein Atlas. URL: <https://www.proteinatlas.org/ENSG00000169228-RAB24/pathology>

Feng, Y., D. He, Z. Yao, and D.J. Klionsky. 2014. The machinery of macroautophagy. *Cell Res*. 24:24–41. doi:10.1038/cr.2013.168.

Flieswasser, T., V. Camara-Clayette, A. Danu, J. Bosq, V. Ribrag, P. Zabrocki, L. Van Rompaey, H. de Haard, K. Zwaenepoel, E. Smits, P. Pauwels, and J. Jacobs. 2019. Screening a Broad Range of Solid and Haematological Tumour Types for CD70 Expression Using a Uniform IHC Methodology as Potential Patient Stratification Method. *Cancers*. 11:1611. doi:10.3390/cancers11101611.

GitHub - maweigert/neubias\_academy\_stardist. URL: [https://github.com/maweigert/neubias\\_academy\\_stardist](https://github.com/maweigert/neubias_academy_stardist)

He, H., F. Dai, L. Yu, X. She, Y. Zhao, J. Jiang, X. Chen, and S. Zhao. 2002. Identification and Characterization of Nine Novel Human Small GTPases Showing Variable Expressions in Liver Cancer Tissues. *Gene Expr.* 10:231–242. doi:10.3727/000000002783992406.

Humphries, M.P., P. Maxwell, and M. Salto-Tellez. 2021. QuPath: The global impact of an open source digital pathology system. *Comput. Struct. Biotechnol. J.* 19:852–859. doi:10.1016/j.csbj.2021.01.022.

Hutagalung, A.H., and P.J. Novick. 2011. Role of Rab GTPases in Membrane Traffic and Cell Physiology. *Physiol. Rev.* 91:119–149. doi:10.1152/physrev.00059.2009.

Klionsky, D.J. 2008. Autophagy revisited: A conversation with Christian de Duve. *Autophagy.* 4:740–743. doi:10.4161/auto.6398.

Latest qupath topics - Image.sc Forum. URL: <https://forum.image.sc/tag/qupath>

Liu, W., J.-M. Yi, Y. Liu, C. Chen, K.-X. Zhang, C. Zhou, H.-E. Zhan, L. Zhao, S. Morales, X.-L. Zhao, and H. Zeng. 2021. CDK6 Is a Potential Prognostic Biomarker in Acute Myeloid Leukemia. *Front. Genet.* 11:600227. doi:10.3389/fgene.2020.600227.

Marinković, M., M. Šprung, M. Buljubašić, and I. Novak. 2018. Autophagy Modulation in Cancer: Current Knowledge on Action and Therapy. *Oxid. Med. Cell. Longev.* 2018:1–18. doi:10.1155/2018/8023821.

Mazières, J., W. Brugger, F. Cappuzzo, P. Middel, A. Frosch, I. Bara, G. Klingelschmitt, and B. Klughammer. 2013. Evaluation of EGFR protein expression by immunohistochemistry using H-score and the magnification rule: Re-analysis of the SATURN study. *Lung Cancer.* 82:231–237. doi:10.1016/j.lungcan.2013.07.016.

Mizushima, N., Y. Ohsumi, and T. Yoshimori. 2002. Autophagosome formation in mammalian cells. *Cell Struct. Funct.* 27:421–429. doi:10.1247/csf.27.421.

Olkkonen, V.M., P. Dupree, I. Killisch, A. Lütcke, M. Zerial, and K. Simons. 1993. Molecular cloning and subcellular localization of three GTP-binding proteins of the rab subfamily. *J. Cell Sci.* 106 ( Pt 4):1249–1261.

Parzych, K.R., and D.J. Klionsky. 2014. An Overview of Autophagy: Morphology, Mechanism, and Regulation. *Antioxid. Redox Signal.* 20:460–473. doi:10.1089/ars.2013.5371.

Projects — QuPath 0.2.3 documentation. URL: <https://qupath.readthedocs.io/en/latest/docs/tutorials/projects.html>

Qin, F., J. Zhang, J. Gong, and W. Zhang. 2021. Identification and Validation of a Prognostic Model Based on Three Autophagy-Related Genes in Hepatocellular Carcinoma. *BioMed Res. Int.* 2021:1–15. doi:10.1155/2021/5564040.

QuPath | Quantitative Pathology & Bioimage Analysis. URL: <https://qupath.github.io/>

QuPath H-Scoring - Image Analysis - Image.sc Forum. URL: <https://forum.image.sc/t/qupath-h-scoring/28483>

RAB24 Antibody 11445-1-AP | Proteintech. URL: <https://www.ptglab.com/products/RAB24-Antibody-11445-1-AP.htm>

Russell, R.C., H.-X. Yuan, and K.-L. Guan. 2014. Autophagy regulation by nutrient signaling. *Cell Res.* 24:42–57. doi:10.1038/cr.2013.166.

Schmidt, U., M. Weigert, C. Broaddus, and G. Myers. 2018. Cell Detection with Star-convex Polygons. *ArXiv180603535 Cs.* 11071:265–273. doi:10.1007/978-3-030-00934-2\_30.

Seitz, S., Y. Kwon, G. Hartleben, J. Jülg, R. Sekar, N. Krahmer, B. Najafi, A. Loft, S. Gancheva, K. Stemmer, A. Feuchtinger, M. Hrabe de Angelis, T.D. Müller, M. Mann, M. Blüher, M. Roden, M. Berriel Diaz, C. Behrends, J. Gilleron, S. Herzig, and A. Zeigerer. 2019. Hepatic Rab24 controls blood glucose homeostasis via improving mitochondrial plasticity. *Nat. Metab.* 1:1009–1026. doi:10.1038/s42255-019-0124-x.

Separating stains — QuPath 0.2.3 documentation. URL: [https://qupath.readthedocs.io/en/latest/docs/tutorials/separating\\_stains.html](https://qupath.readthedocs.io/en/latest/docs/tutorials/separating_stains.html)

Wang, S., C. Hu, F. Wu, and S. He. 2017. Rab25 GTPase: Functional roles in cancer. *Oncotarget*. 8:64591–64599. doi:10.18632/oncotarget.19571.

Welcome to QuPath! — QuPath 0.2.3 documentation. URL: <https://qupath.readthedocs.io/en/latest/>

Yamamuro, T., T. Kawabata, A. Fukuhara, S. Saita, S. Nakamura, H. Takeshita, M. Fujiwara, Y. Enokidani, G. Yoshida, K. Tabata, M. Hamasaki, A. Kuma, K. Yamamoto, I. Shimomura, and T. Yoshimori. 2020. Age-dependent loss of adipose Rubicon promotes metabolic disorders via excess autophagy. *Nat. Commun.* 11:4150. doi:10.1038/s41467-020-17985-w.

Ylä-Anttila, P., and E.-L. Eskelinen. 2018. Roles for RAB24 in autophagy and disease. *Small GTPases*. 9:57–65. doi:10.1080/21541248.2017.1317699.

Ylä-Anttila, P., E. Mikkonen, K.E. Happonen, P. Holland, T. Ueno, A. Simonsen, and E.-L. Eskelinen. 2015. RAB24 facilitates clearance of autophagic compartments during basal conditions. *Autophagy*. 11:1833–1848. doi:10.1080/15548627.2015.1086522.

Yun, C., and S. Lee. 2018. The Roles of Autophagy in Cancer. *Int. J. Mol. Sci.* 19:3466. doi:10.3390/ijms19113466.