



Article Dynamic NIR Fluorescence Imaging and Machine Learning Framework for Stratifying High vs. Low Notch-Dll4 Expressing Host Microenvironment in Triple-Negative Breast Cancer

Shayan Shafiee ¹, Jaidip Jagtap ², Mykhaylo Zayats ³, Jonathan Epperlein ³, Anjishnu Banerjee ⁴, Aron Geurts ⁵, Michael Flister ⁵, Sergiy Zhuk ³ and Amit Joshi ^{1,*}

- ¹ Department of Biomedical Engineering, Marquette University and Medical College of Wisconsin, Milwaukee, WI 53226, USA
- ² Department of Radiology, Mayo Clinic, Rochester, MN 55905, USA
- ³ IBM Research Europe, D15 HN66 Dublin, Ireland
- ⁴ Division of Biostatistics, Medical College of Wisconsin, Milwaukee, WI 53226, USA
- ⁵ Department of Physiology, Medical College of Wisconsin, Milwaukee, WI 53226, USA
- * Correspondence: ajoshi@mcw.edu

Simple Summary: Breast cancer is a disease that is affected by both the tumor cells and the host environment. It is well known that the tumor blood vessels are aberrant in structure and function due to rapid angiogenesis, and this aberrant vasculature plays a major role in drug delivery and therapy response of breast cancer. Dll4 is a protein that helps control the growth of blood vessels in tumors. This study used near-infrared optical imaging and a novel machine learning framework to determine if Dll4 levels can be predicted from simple noninvasive imaging assays. The eventual results of this study may help physicians decide if a given triple-negative breast cancer patient will benefit from a Dll4 targeted therapy.

Abstract: Delta like canonical notch ligand 4 (Dll4) expression levels in tumors are known to affect the efficacy of cancer therapies. This study aimed to develop a model to predict Dll4 expression levels in tumors using dynamic enhanced near-infrared (NIR) imaging with indocyanine green (ICG). Two rat-based consomic xenograft (CXM) strains of breast cancer with different Dll4 expression levels and eight congenic xenograft strains were studied. Principal component analysis (PCA) was used to visualize and segment tumors, and modified PCA techniques identified and analyzed tumor and normal regions of interest (ROIs). The average NIR intensity for each ROI was calculated from pixel brightness at each time interval, yielding easily interpretable features including the slope of initial ICG uptake, time to peak perfusion, and rate of ICG intensity change after reaching half-maximum intensity. Machine learning algorithms were applied to select discriminative features for classification, and model performance was evaluated with a confusion matrix, receiver operating characteristic curve, and area under the curve. The selected machine learning methods accurately identified host Dll4 expression alterations with sensitivity and specificity above 90%. This may enable stratification of patients for Dll4 targeted therapies. NIR imaging with ICG can noninvasively assess Dll4 expression levels in tumors and aid in effective decision making for cancer therapy.

Keywords: cancer; tumor microenvironment modifier; notch-DLL4; consomic xenograft model; machine learning; binary classification; dynamic enhanced NIR imaging; indocyanine green; time series; tumor detection

1. Introduction

Breast cancer heterogeneity has been extensively studied and it has enabled the classification and categorization of tumors into molecular subtypes depending on the overexpression of antigens or hormone receptors on tumor cells [1]. Identification of tumor



Citation: Shafiee, S.; Jagtap, J.; Zayats, M.; Epperlein, J.; Banerjee, A.; Geurts, A.; Flister, M.; Zhuk, S.; Joshi, A. Dynamic NIR Fluorescence Imaging and Machine Learning Framework for Stratifying High vs. Low Notch-Dll4 Expressing Host Microenvironment in Triple-Negative Breast Cancer. *Cancers* **2023**, *15*, 1460. https://doi.org/10.3390/ cancers15051460

Academic Editor: Mark W. Dewhirst

Received: 30 December 2022 Revised: 22 February 2023 Accepted: 23 February 2023 Published: 25 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). subtypes improves cancer patients' care and prognosis by tailoring therapies to the subtypes [2–4]. Breast and many other cancers are highly heritable, yet most causative variants are unknown, and most of the known risk variants are considered tumor-cell-autonomous, with far less emphasis placed on identifying the role of germline variants impacting the tumor microenvironment (TME). The TME is a complex and dynamic system that includes cancer cells, stromal cells, blood vessels, and extracellular matrix [5] and plays a significant role both in tumor cell proliferation and in chemo- or radiotherapy delivery and response [5–7]. There is growing evidence that heritable modifiers of the tumor microenvironment can profoundly impact tumor behavior and response to diagnostic and therapeutic interventions [8–12].

Tumor blood vessels have abnormal structure and function, which leads to heterogeneity in blood perfusion both temporally and spatially [13]. This heterogeneity has multiple adverse consequences, including limiting the access of blood-borne drugs and effector immune cells to poorly perfused regions of tumors [14]. As a result, these areas become hypoxic and have low extracellular pH [15]. Hypoxia has been shown to play a significant role in tumor progression and metastasis by inducing genetic instability, angiogenesis, immunosuppression, inflammation, and resistance to cell death by apoptosis and autophagy [16,17]. Anti-angiogenic drugs are designed to target the vasculature in order to starve tumors and prevent them from growing. However, recent studies have shown that the efficacy of these drugs may be limited by specific biomarkers and pathways associated with resistance [15]. For example, it has been shown that some patients may not benefit from anti-VEGF therapies if they have elevated levels of plasma sVEGFR1 [18]. Similar outcomes have been observed with increased levels of SDF1 α and anti-VEGF therapies [19]. Further the vascular TME and therapy response differs in primary tumor and metastasis sites [20] and the anatomic location [21–23]. Thus, characterizing angiogenesis in tumors holistically may have therapeutic implications.

The process of tumor angiogenesis is closely regulated by a balance between promoting and suppressing angiogenic factors [24,25]. Delta like canonical notch ligand 4 (Dll4) is a protein-coding gene that provides instructions for making a protein part of a signaling pathway known as the notch pathway, which is essential for the normal development of many tissues throughout the body, affecting cell functions [26,27], modulating tumor angiogenesis [28], promoting vessel maturation, and inhibiting vessel sprouting by inducing apoptosis of tip endothelial cells (TECs) [28–30]. Dll4 is overexpressed in various types of cancer, including breast, ovarian, and colorectal cancer, and has been shown to promote tumor angiogenesis, growth, and metastasis by interacting with receptors on endothelial cells (ECs) [31–34]. Blockade of Dll4 activity results in enhanced vessel sprouting and increased vascular permeability [29,30,35], but anti-Dll4 therapy has not been universally successful, as Dll4 has been shown to have both pro-tumorigenic and anti-tumorigenic effects depending on the context of its expression [34,36,37]. We recently reported that Dll4 expression on the host TME rather than on tumor cells determines the EPR or enhanced permeation and retention effect in breast tumor xenografts and thus governs nanomedicine delivery and therapy response [38].

Despite the increasing evidence about the function of germline genetic modifiers, such as Dll4, in TME heterogeneity and enhanced permeability and retention (EPR) effects, the underlying influencers have mainly remained unexplored because of the lack of a systematic approach to studying them. Therefore, we developed the Consomic Xenograft Model (CXM) as a strategy for mapping heritable modifiers of TME heterogeneity. In the CXM, human breast cancer cells are orthotopically implanted into consomic xenograft host strains. These strains are derived from two parental strains with different susceptibilities to breast cancer. Salt-sensitive (SS) rats were employed as a tumor promoting strain, while Brown Norway (BN) rats were used as a tumor suppressing strain. A sequence of consomic strains were generated with chromosomes of SS rats replaced by those of BN rats one at a time and used for breast tumor xenograft studies [7,38,39]. Because the host backgrounds genetically differ by one chromosome, whereas the tumor cells are unvaried, any observed phenotypic

differences are due to TME modifier(s) and can be linked to a single chromosome. These modifiers can be further localized by congenic mapping (inbred strains containing a given sub-chromosomal region in their genome). By combining CXM with dynamic epifluorescence near-infrared (DE-NIR) imaging, systemic injections of indocyanine green (ICG) through a tail vein in tumor-bearing rats, and multiparametric analysis of pharmacokinetic modeling, we localized and identified the function of the vascular-specific Dll4 allele on rat chromosome 3 (RNO3) as a heritable host TME modifier of EPR [38]. The SS.BN^{3IL2Rγ-} CXM strain with low-level expression of Dll4 (referred to as Dll4-) had significant tumor growth inhibition compared with the parental SS^{IL2Rγ-} strain with higher expression of Dll4 (Dll4+), despite a paradoxical increase in tumor blood vessel density in Dll4+. Further analysis revealed that the changes in the Dll4+ tumors were accompanied by altered expression of Dll4, which was previously linked with nonproductive angiogenesis. Additionally, Dll4 was found to be co-localized within a host TME modifier locus (Chr3: 95–131 Mb) identified by congenic mapping and correlated with the phenotypic differences observed at the consomic level [7,39,40].

The inheritance of functionally different Dll4 alleles can influence the efficacy of nanoparticle (NP) therapy, and previous results indicate that inherited microvascular distribution patterns, rather than overall NP uptake, ultimately determine the effectiveness of NP-mediated photothermal therapy (PTT). Consequently, patients with high endothelial Dll4 expression can be selected for treatment with anti-Dll4 targeted nanoparticles as opposed to patients with low Dll4 expression, where PEGylated nanoparticles will provide sufficient therapy response [38]. Recent advances in dynamic vascular imaging techniques, such as DCE-MRI and perfusion computed tomography, have facilitated the investigation of the time kinetics of a contrast agent to extract multiple vascular parameters and have been successfully applied in clinical trials of anti-angiogenic drugs [41,42]. However, these techniques have certain drawbacks, including a lack of high temporal resolution and the need for a heavy hardware system with sophisticated analysis software. Dynamic NIR fluorescence imaging, on the other hand, offers a sufficient and effective alternative to other dynamic vascular imaging techniques for characterizing germline-dependent vascular phenotypes [7,43]. This has led to the combination of these modalities, such as in the paired agent MRI-coupled fluorescence tomography approach for noninvasive quantification of paired-agent uptake in response to anti-angiogenesis therapy in vivo [44].

As the field of artificial intelligence continues to advance, researchers are increasingly utilizing AI techniques, particularly machine learning, to develop predictive models that can support effective decision making in various domains including cancer therapy selection [45,46]. Previous research has investigated the use of machine learning algorithms to analyze near-infrared (NIR) signal intensity and perfusion patterns to differentiate between healthy, benign, and malignant tissue [47]. This work demonstrated that the signal intensity time course of an FDA-cleared near-infrared dye ICG inflow during the wash-in phase and ICG outflow during the wash-out phase could serve as significant markers for tissue distinction. This finding offers a new method for noninvasive tissue distinction and has prognostic potential [48,49]. However, there remains a need for further exploration of the use of machine learning for classifying host genetic tumor microenvironment (TME) modifiers and predicting therapy responses based on dynamic contrast-enhanced imaging of tumors, particularly DE-NIR fluorescence imaging data [47].

2. Hypothesis and Objective

We hypothesize that the observation of subtle differences in vasculature structure and perfusion patterns characterized by ICG inflow and outflow using DE-NIR imaging could be used to differentiate between inherited tumor vascular microenvironment differences, such as Dll4 expression levels. We propose an experimental framework to noninvasively assess Dll4 expression levels in tumors based on the NIR signal intensity time course of perfusion patterns characterized by ICG time kinetics to develop a predictive model to support effective decision making in cancer therapy. Herein, we used two rat-based CXM strains of breast cancer, $SS^{IL2R\gamma-}(Dll4+)$ and $SS.BN3^{IL2R\gamma-}(Dll4-)$ [7,38,50], as well as eight congenic xenograft strains, CG1–CG8 (Figure 1a,b), to assess the impact of germline TME vascular heterogeneity on the signal intensity of DE-NIR imaging with systemically delivered ICG. Principal component analysis (PCA)-based decomposition of time-dependent epifluorescence videos (image stacks) was used for visualization and anatomical segmentation of tumors, liver, lungs, and fat pads [7]. In addition, we utilized modified principal component analysis (PCA)-based anatomical segmentation techniques to identify and analyze regions of interest (ROIs) representing potential tumors within the current dataset. To gather further information, we calculated the average NIR intensity for each ROI by analyzing the brightness of individual pixels at each time interval, resulting in a series of intensity measurements for each ROI. From this analysis, several easily interpretable features were extracted, including the slope of the initial uptake of ICG, the time it takes to reach peak perfusion, and the rate of ICG intensity changes once the half-maximum intensity is reached (which, to the best of our knowledge, has not been previously reported in the literature). We then applied a subset of machine learning algorithms, including Support Vector Machines (SVMs), Naive Bayesian Classifiers (NBCs), Generalized Additive Models (GAMs), Decision Trees (DTs), Nearest Neighbors (NN), and Logistic Regression (LR) to select the most discriminative features for classification. The performance of the model was evaluated using confusion matrix, receiver operating characteristic curve (ROC), and the area under the curve. To further evaluate our hypothesis of detecting Dll4 expression levels from DE-NIR imaging and test the generalizability of our framework, we conducted a secondary performance evaluation method using congenic groups with high and low Dll4 expression levels. The classification models were trained based on the selected features, and the performance of the model was tested on the remaining congenic groups. We demonstrate that robust ML methods can identify the alterations in host Dll4 expression from the tumor dynamic imaging datasets, and thus these methods can potentially stratify patients for Dll4 targeted therapies.





(blue). (b) Schematic representing congenic strains and vascular-specific Dll4 protein expression, generated by introgressing BN(Dll4–) chromosome 3 (blue) into the SS^{IL2R γ -} strain (red) via marker-assisted breeding. Thin black bars represent confidence intervals of chromosomal regions that could be BN or SS. The horizontal black line indicates the location of vascular-specific Dll4 protein on chromosome. The dotted lines represent the location of SSLP markers. (c) Transgenically labeled human MDA-MB-231 triple-negative breast cancer cells were orthotopically implanted in the mammary fat pad of parental, consomic, and congenic rats. (d) Schematic diagram of the dynamic epifluorescence NIR imaging setup. A bifurcated optical fiber bundle delivers 785 nm excitation light for uniform illumination of the rat body surface. A 16-bit deep-cooled intensified charge-coupling device camera was used to image the rats through computer-controlled software.

3. Materials and Methods

All methods have been carried out in accordance with relevant guidelines and regulations. Approved protocols by the Medical College of Wisconsin Institutional Biosafety Committee (IBC) and Institutional Animal Care and Use Committee (IACUC) were followed. All live animal experiments are reported per the ARRIVE guidelines' recommendations [51]. All results were rigorously adjusted for multiple comparisons.

3.1. Animals

All animal protocols employed in this study were approved by the Institutional Animal Care and Use Committee (IACUC), Medical College of Wisconsin (MCW). The MCW has an Animal Welfare Assurance (Assurance number D16-00064 (A3102-01)) on file with the Office of Laboratory Animal Welfare, National Institutes of Health (NIH). Animal experiments were performed according to the relevant guidelines and regulations and in compliance with the Guide for the Care and Use of Laboratory Animals published by the US National Institutes of Health (NIH publication no.85-23, revised 1996). SS.BN3 rats were developed as part of the Consomic Xenograft Model at the Medical College of Wisconsin (MCW) [40]. SS and SS.BN3 rats were purchased from the Rat Research Models Service Center at the Medical College of Wisconsin [52]. All rats were provided reverse osmosis hyper-chlorinated water ad libitum. All animal experiments were performed on anesthetized animals. The animal was placed in a transparent induction chamber to induce anesthesia. Isoflurane was delivered through a precision vaporizer and compressed O_2 to the chamber. For induction, the percentage of isoflurane was up to 5%. Once the animal was unconscious, it was removed from the chamber. The unconscious animal was then placed on a warm surface and fitted with a nose cone attached to the vaporizer in the presence of a scavenging system and oxygen source. At this point, the concentration of isoflurane was reduced to this level that maintained the correct anesthesia plane, usually between 0.5 and 3%. After the end of the experiment, or when other criteria for animal protocols were justified, rats were euthanized. Rats were placed in an approved euthanasia chamber and exposed to CO₂ from a compressed gas cylinder until the animal was no longer breathing. To ensure death in rats, a pneumothorax was created via thoracotomy for rats weighing more than 200 g. For rats weighing less than 200 g, a pneumothorax was created, or a cervical dislocation was performed.

As previously described [39,40], consomic strains (SS and SS.BN3 rats) were generated by sequentially replacing SS chromosomes with the outbred wild-type and tumor-resistant strain of Brown Norway (BN) rats referred to as SS.BN#, which are reported for their tumorigenic potential, where # refers to the chromosome number. These parental SS and consomic SS.BN# strains were genetically ablated by knocking down the IL2R γ gene to allow the grafting and growth of human cancer cell lines. Such immunocompromised strains are labeled as SS^{IL2R γ -(Dll4+) and SS.BN#^{IL2R γ -(Dll4-).}}

Previous research has localized inherited modifier(s) of TME vascular heterogeneity to RNO3 by CXM mapping and further narrowed by congenic mapping to a 36 Mb locus containing Dll4 alleles with distinct vascular expression patterns in the SS.BN3^{IL2R γ - consomic (Dll4–) and SS^{IL2R γ - (Dll4+) rat strains, and verified via species-specific RNA}}

sequencing and immunohistochemistry that strains inheriting the SS Dll4 allele on chromosome 3 have higher Dll4 expression on tumor-associated endothelium and that the blood vessel tortuosity and dysfunction increased in Dll4- strains [38,39]. Since there are many other candidate alleles on chromosome 3 that could also potentially account for the observed differences in therapeutic efficacy between the SS.BN3^{IL2R γ -} and SS^{IL2R γ -} strains [38], to further investigate the potential contribution of Dll4 to inherited tumor vascular heterogeneity, eight novel SS.BN3^{IL2R γ - congenic xenograft host strains (CG1 to} CG8) were constructed by introgression of the F1 progeny and F2 generation to capture different regions of RNO3 by marker-assisted selection, as described previously [50,53]. The exclusion congenic mapping localized a 7.9 Mb candidate region (marked by the SSLP marker D3Mgh11) that was associated with inherited tumor vascular heterogeneity and contained the Dll4 locus. As a result, eight new congenic strains CGN(s-e) were generated, where N (1 to 8) refers to the congenic group, while s and e refer to the starting and ending of Simple Sequence Length Polymorphism (SSLP) marker regions, respectively. This resulted in generating CG1(D3Rat26-D3Mgh30), CG2(D3Rat222-D3Got42), CG3(D3Rat222-D3Mco33), CG4(D3Rat164-D3Rat218), CG5(D3Rat26-D3Mco218), CG6(D3Rat86-D3Rat218), CG7(D3Mgh13-D3Rat218), and CG8(D3Rat160-D3Rat218) congenic groups (Figure 1a,b).

3.2. Cell Culture and Triple-Negative Breast Cancer Xenografts

As previously described [38], triple-negative breast cancer MDA-MB-231 cells were maintained in DMEM media (Sigma, Burlington, MA, USA) supplemented with 10% FBS (Gibco, New York, NY, USA) and 1% penicillin and streptomycin (Lonza, Cohasset, MN, USA) and incubated in 5% CO₂ at 37 °C. These cells (6×10^6) in 50% Matrigel were orthotopically implanted into the mammary fat pads (MFP) of 4- to 6-week-old female Dll4+ (n = 8) and Dll4– (n = 17) rats and eight congenic strains CG1 (n = 19), CG2 (n = 2), CG3 (n = 26), CG4 (n = 12), CG5 (n = 28), CG6 (n = 12), CG7 (n = 5), and CG8 (n = 4) (Figure 1c) [54]. Tumors were treated after 10 days of implantation at an approximate size of 600 mm³, consistent across all rat strains.

3.3. In Vivo NIR Fluorescence Imaging

A customized NIR imaging system was assembled for imaging the rats. A bifurcated optical fiber bundle was used to deliver 785 nm excitation light (~5 mW/cm² power at the surface, diode laser, ThorLabs Inc., Newton, NJ, USA) from two positions for uniform illumination of the entire rat body surface. A 16-bit deep-cooled intensified charge-coupled device camera (PIMAX4 ICCD, Princeton Instruments, Trenton, NJ, USA) equipped with 830 nm long-pass filter positioned following a holographic notch rejection filter in the optical path (ThorLabs Inc.) was used to image the rats through computer-controlled LightField[®] software (Teledyne Princeton Instruments, Trenton, NJ, USA) (Figure 1d).

Dynamic contrast-enhanced NIR fluorescence imaging was performed on anesthetized rats, as reported previously for 800 nm NIR imaging [7]. In this study, the setup was used for imaging the whole body. A total of 133 rats (Dll4+ (n = 8), Dll4– (n = 17)) and eight congenic strains CG1 (n = 19), CG2 (n = 2), CG3 (n = 26), CG4 (n = 12), CG5 (n = 28), CG6 (n = 12), CG7 (n = 5), and CG8 (n = 4) were imaged. NIR imaging was performed for approximately 6 min following ICG injection with the CCD array hardware binned to 256 × 256 with a frame rate of 10.6 fps. A total of 3000 frames were captured for each imaging case, including about 50 frames for background correction. ICG (MP Biomedicals) was delivered in an intravenous bolus of 1200 μ M ICG/200 g body weight into the tail vein via a catheter with a 32-gauge needle tip connected to a syringe pump (Harvard Apparatus PHD 2000 syringe pump, Holliston, MA, USA) operated at a speed of 0.2 mL/s. The injected volume was calibrated to provide a body-weight-equilibrated dose to each rat.

3.4. Denoising and Motion Correction

Image processing and data analysis were performed in MATLAB (R2021b MATH-WORKS Inc.) software. The time-dependent image frames were assembled as 3D arrays

(two spatial and one temporal dimensions) for all animals. A custom-designed breathing correction method with a low-pass temporal filter combined with a 1D wavelet-based denoising was used to filter the high-frequency jitter generated by the animal's respiratory motion from the fluorescence kinetic sequences of each pixel, as described previously [7]. An average of pre-ICG injection frames (acquired in the ~5 s before ICG injection) was used as background, incorporating contributions from CCD noise and excitation light leakage from emission filters and subtracted from all the frames. (Refer to Video S1 for respiratory motion corrected time course images of ICG biodistribution).

3.5. Principal Component Analysis for Extraction of Spatial Patterns of Internal Organs

First, motion correction and background subtraction were performed on the imaging data described in the previous steps. This was carried out to remove any potential artifacts or noise that could affect the accuracy of the principal component analysis. Next, the data were decomposed by PCA along the time dimension using MATLAB software following the previously published methods [55,56]. This resulted in converting the imaging data to a k-component vector for each pixel, where k is the number of time-frames in the original dataset. The PCA on the dynamic fluorescence image was used to extract spatial patterns of internal organs linked to statistically similar kinetic behaviors. This was carried out by comparing the k-component vectors for each pixel and identifying those that displayed similar patterns over time. The contribution of the first six principal components on a time basis is illustrated in Figure S1.

3.6. PCA Ranking Tumor Detection

The ROI detection module employed three steps in order to identify the tumor area in images (Figure 2a): (1) spatial alignment, (2) PCA ranking and selection, and (3) ROI selection and masking. First, images were registered to a reference image using a rigid body transformation in order to ensure consistent spatial alignment for the detection of the tumor area in subsequent steps. The variability in the visual appearance of internal organs or tumors within certain principal components necessitated the adaptation of existing methods for ranking normalized components based on their two-dimensional cross-correlation (2DCC) with a reference image containing the tumor. This was achieved by generating a stack of the first ten normalized principal components and applying a two-dimensional cross-correlation function (xcorr2 in MATLAB) to each component. The principal components were ranked according to the correlation scores obtained and a semi-automatic algorithm was then utilized to carry out the subsequent two steps for tumor ROI generation.

To select the appropriate principal component for tumor identification, the algorithm prioritized the PC with the highest likelihood of containing the tumor tissue based on the ranking from the previous step. Once the proper PC was selected, we used 2DCC to estimate the tumor's location in the frame. This was performed by creating a stack of four reference images using PCs that clearly visualized the tumor and generated a bounding box around the tumor. Next, we applied basic morphological dilation and image arithmetic operations to emphasize the tumor's boundaries. A threshold was then applied to separate the tumor from the background. Finally, we used Blob Analysis (Computer Vision Toolbox[™], MATLAB, MathWorks, MA, USA) to extract the centroid and exact location of the tumor in the frame, generating a region of interest (ROI). The use of a graphic user-interface-enabled semi-automatic platform allowed for real-time evaluation and adjustment of the algorithm's performance if necessary.



Figure 2. (a) Diagram of the process for detecting tumors and generating regions of interest (ROIs). The ROI detection module consists of three steps: spatial alignment, PCA ranking and selection, and ROI selection and masking. First, NIR time series data are registered to a reference image using a rigid body transformation. Next, PC reconstructions are ranked based on their 2D cross-correlation with a reference image featuring a clearly visible tumor. The tumor's location and a bounding box (shown in yellow) are estimated based on the 2D cross-correlation score of the selected PC and reference images featuring different tumors. The tumor's boundaries are enhanced and distinguished from the background through morphological dilation and thresholding. The resulting ROI is used to mask the NIR time series data. (b) Plot of the mean intensity of ICG biodistribution fluorescence kinetic data over 300 s for the tumor (red) and four fat pads (blue), normalized to a range of 0 to 1.

3.7. Video Processing

The underlying assumption for the NIR fluorescence intensity video analysis is that intensity was proportional to the ICG concentration in the tissue, which has been shown in the literature [57].

From each available video, p, between 4 and 5 fluorescence time series are extracted, one based on the tumor ROI generated from earlier steps, and between three and four from mammary fat pads (MFPs) (Figure 2b). For each ROI, r, and each included time step, t, the mean brightness of the pixels inside that ROI in the NIR video is stored as $I_{p,r}(t)$. The result is a collection of time series $I_{p,r}(t)$, where p ranges from 1 to N (the number of animals in each group in our dataset), r ranges from 4 to the number of ROIs generated in the specific animal (up to 5, 1 tumor and 3 or 4 fat pads), and t from 0 to 3000 (Figure 2b).

3.8. Peak and Latency Estimation

First, we started with smoothing the data to exclude any potential noise from motion artifacts. The smoothing was conducted using a Savitzky–Golay filter of order 3 with window length 31 [58]. The parameters for this filter were determined through manual annotation of peak and latency in a subset of the data, with the goal of minimizing average estimation error. Subsequently, we used a MATLAB function to identify the time point of maximum intensity, which corresponded to the peak of the fluorescence signal. We employed a custom latency detector script that analyzed the smoothed derivative of the curves and identified the first "robust" zero crossing as the latency point ($L_{p,r}(t)$) [48].

The features were chosen in two steps as previously described (Figure 3) [47,48]: first, the following characteristic numbers were chosen for each normalized time series individually.



Abbreviation	Feature	Description
L	Latency	The point in time when fluorescence starts rising beyond the initial background
		fluorescence
T _{1/2}	Time to Half Peak Intensity	The time difference between the half peak intensity time point and the latency
TTP	Time to Peak	The time difference between the peak intensity time point and the latency
U	Upslope	The average slope between initial ICG arrival(latency) and the peak
Ds	Downslopes	The downslope between the peak and t seconds further.
HIFs	Half Intensity Forward slopes	The slope between the half peak Intensity and t seconds further.

Figure 3. Feature extraction. The normalized mean intensity of ICG taken over ROIs is plotted for the first 75 s (blue). The time series was smoothed using a Savitzky–Golay filter of order 3 with window length 31 (red). The insets show enlarged curve regions to visualize the characteristics and relationships of the features. S = $\{2, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 23, 25, 30, 35, 40, 50, 60, 70, 80\}$.

The time to peak $(TTP_{p,r})$ is simply the time difference between the peak intensity time point and the latency $L_{p,r}$.

The upslope $(U_{p,r})$ is computed as

$$U_{p,r} = 1 - \frac{I_{p,r}(L_{p,r})}{TTP_{p,r}},$$
(1)

which is the average slope between initial ICG arrival and the peak. The downslopes (*DS*) are the downslopes between the peak and *S* seconds further:

$$D_{p,r} = 1 - \frac{I_{p,r}(L_{p,r} + TTP_{p,r} + S)}{S},$$

$$S \in \{2, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 23, 25, 30, 35, 40, 50, 60, 70, 80\}.$$
(2)

The time ratio (TR) is the ratio between TTP_p and when $I_p(t)$ reaches half the peak values. The half intensity forward (HIF) is the average slopes between when $I_{p,r}(t)$ reaches half the peak values $(T_{1/2p,r})$ and *S* seconds further, so

$$HIF_{p,r} = 1 - \frac{I_{p,r}(L_{p,r} + T_{1/2p,r} + S)}{I},$$

$$S \in \{2, 4, 5, 6, 8, 10, 12, 14, 16, 18, 20, 23, 25, 30, 35, 40, 50, 60, 70, 80\}.$$
(3)

$$DS_{p,r,avg} = median \ s \in Window \ D(S+s)_{p,r'}$$
(4)

$$HIF_{p,r,avg} = median \ s \in Window \ D(S+s)_{p,r'}$$
(5)

$$Window = \{-1.5, -1.3, \dots, 1.3, 1.5\}$$
 seconds.

These features provide insight into the tumor's ICG uptake and decay. *U* and *TTP* relate to the initial uptake of ICG, while *DS* relates to the decay of ICG fluorescence [48]. *TR* is a measure of the temporal inhomogeneity of the initial uptake, and *HIF* is a measure of the temporal inhomogeneity of both the initial uptake and decay of ICG fluorescence.

To address inter-animal variation, we propose a feature design that relates the features of tumor intensity ($F_{r=tumor}$) to the features of healthy tissue intensity ($F_{r=fatpad}$) in the same animal. The median value of a feature across the healthy tissue (fat pad) is chosen as a reference value, and each feature is defined as its percentage difference to that reference value. This results in a single normalized feature for each animal. By using the median as an average that is robust to outliers, our feature design allows for a more accurate comparison of tumor intensity across different animals.

$$F_n = \frac{F_{tumor} - median(F_{fat pad})}{median(F_{fat pad})},$$
(6)

3.10. Classification Algorithms and Feature Selection

The feature extraction and design required specialized knowledge of the tumor microenvironment (TME) and its impact on near-infrared intensity. However, the classification based on the inheritances of Dll4 can be considered a standard binary classification problem with feature selection as a sub-problem: given the features of an animal for which the Dll4 inheritance is unknown, we want to assign the label "Dll4high" or "Dll4low" to it. We also want to investigate which small subset of features performs best [44]. We restricted ourselves to a subset of available ML algorithms which were reported to perform best with intensity time series classification. [48], We excluded neural networks from consideration and evaluated Support Vector Machines (SVMs), Naive Bayesian Classifiers (NBs), Generalized Additive Models (GAMs), Decision Trees (DTs), Nearest Neighbors (NN), and Logistic Regression (LR).

3.11. Primary Classification

DTs and SVMs using the full set of 86 features were trained. Each ML model was tuned to the training set in an internal cross-validation procedure of 10-fold. This process was repeated 20 times, and the performance metrics are reported as classifier performance. One of the key metrics used to evaluate the performance of a classification algorithm is accuracy, which measures the proportion of correctly classified instances in the dataset. Other important metrics include sensitivity, which measures the proportion of negative instances that were correctly classified, and specificity, which measures the proportion of negative instances that were correctly classified. In our case, we are interested in classifying animals as either Dll4high or Dll4low, and we use the average metric (A_{score}) for each pair of groups to evaluate the performance of the different classification algorithms.

The A_{score} for a given pair of groups is calculated as the average of accuracy, sensitivity, and specificity, as follows:

$$A_{score} \left(Dll4^{high} \right| Dll4^{low} \right) = \frac{Accuracy + Sensitivity + Specificity}{3}, \tag{7}$$

3.12. Congenic Pair Selection

In this study, we describe a two-step method for selecting congenic pairs with high and low Dll4 expression levels for feature selection and hypothesis testing. This method was used to identify well-behaved pairs for binary classification and to identify the pair with the highest classification performance: The primary selection of congenic pairs is based on their primary classification scores (A_{score}) and their performance against parental strains using all available features. The secondary selection is based on the classification performance using the 2 best-performing features.

The congenic pair selection process involved the selection of all possible combinations of congenic pairs with high and low Dll4 expression levels with n > 5, resulting in a total of 12 pairs: Dll4+|Dll4-, Dll4+|CG3-, Dll4+|CG4-, Dll4-|CG1+, Dll4-|CG5+, DII4- | CG6-, CG1+ | CG3-, CG1+ | CG4-, CG3- | CG5+, CG3- | CG6+, CG4- | CG5+, CG4–1CG6+. It should be noted that we only included subgroups with n > 5 in our analysis, to avoid introducing noise into the feature selection process. This is because smaller sample sizes can be more prone to variability and may not represent the larger population [59]. Therefore, we focused on more significant subgroups to ensure our feature selection process was robust and reliable. These 12 pairs have gone through our primary classification algorithm with a split of 75%-25% for testing and training with seeded randomization and proportional distribution of each group in training and testing datasets. Each ML model was tuned to the training set in an internal cross-validation procedure of 10-fold and evaluated by its performance on the test set. This process was repeated 20 times, and the best A_{score} was reported as the metric of classifier performance. The result of this step was used to identify well-behaved congenic pairs for binary classification by calculating the separation score (S_{score}) for each pair of congenic groups with CG#+ and CG#– as below, which is bound between 0 and 1 and reported in Table 1:

$$S_{score} (CG\# + | CG\# -) = \frac{A_{score} (Dll4 + | CG\# -) + A_{score} (CG\# + | Dll4 -) + 2*A_{score} (CG\# + | CG\# -)}{4},$$
(8)

CG– CG+ Dll	i-	
0.74	0.72	0.775
2 0.74	0.75	0.7875
0.83	0.61	0.7925
0.86	0.778	0.8495
2 0.83	0.8	0.845
2 0.86	0.78	0.855
	Q 0.74 2 0.74 9 0.83 9 0.83 2 0.83 2 0.83 2 0.83 2 0.83 2 0.83	Q 0.74 0.72 2 0.74 0.75 9 0.83 0.61 9 0.86 0.778 2 0.83 0.8 2 0.86 0.778 2 0.86 0.78

Table 1. Separation score for group pairs.

The congenic pairs with S_{score} above 80% were selected for secondary congenic pair selection.

For the secondary selection, we evaluated the classification performance (A_{score}) of each pair determined this time by using the two most effective features as described in Section 3.13. We used a 75%–25% split for testing and training, with randomization to ensure a proportionate representation of each group in both datasets.

The pair with the highest classification performance, A_{score} , based on the two most effective features was selected for the final classification model training. To ensure a highquality classification model, we set a minimum threshold for the A_{score} of 0.70 for inclusion in the feature selection process. Congenic groups with an A_{score} below 0.70 were considered to have an insignificant contribution to the classification model and were excluded from further consideration in the feature selection process. This approach allowed us to focus on the most informative feature pairs, improving the overall classification performance of our machine learning model. Each machine learning model was tuned to the training set using a 10-fold cross-validation procedure and evaluated based on its performance on the test set. We repeated this process 20 times and reported the best A_{score} as the classifier's performance.

3.13. Feature Selection and Secondary Classification

We selected the best pair of features in terms of achieved sensitivity, specificity, and accuracy by a two-step procedure as previously reported: DTs and SVMs using the full set of 86 features were trained, and recursive feature elimination (RFE) was performed to refine a much smaller set of best-performing features [60]. RFE is a widely used machine learning classification algorithm that helps in reducing the dimensionality of feature space and selecting a small subset of features that yield the best classification performance. This was achieved through an iterative procedure that uses a ranking criterion to eliminate features one or more at a time. The RFE algorithm started by selecting a subset of features and training a model on this subset. The features were then ranked based on their contribution to the model's performance, and the least important feature was eliminated. The process was then repeated with the remaining features, and the best subset of features was selected based on a model selection criterion [61].

One of the main advantages of RFE is that it helps to reduce the risk of overfitting when the number of features is large, and the number of training patterns is comparatively small [62]. This is because the algorithm selects only a subset of features that are relevant to the classification task, and this helps to avoid the inclusion of irrelevant and redundant features. RFE can be used in conjunction with other techniques such as regularization and support vector machines (SVMs) to further improve the performance of the classification model. In addition, projection methods such as principal component analysis can reduce the feature space's dimensionality before applying RFE [55].

We used a k-fold cross-validation strategy to assess the performance of our model. We also reserved a portion of the training data for primary testing of the model after hyperparameter optimization. Our experiments were conducted with 20 random splits of the training and test datasets, and the mean performance metrics were reported for sensitivity, specificity, and accuracy as A_{score} . To facilitate the interpretation of our results, we limited the number of final features to two. Furthermore, given the small size of our dataset, there was no justification for using high-dimensional feature spaces.

3.14. Data Augmentation

The use of data augmentation has become a popular technique in machine learning and deep learning, especially in the field of computer vision. Data augmentation involves applying random transformations to the training dataset to increase its diversity and improve the performance of a model. In this study, we used data augmentation on raw near-infrared (NIR) image stacks to evaluate the robustness of a classification model. We used a dataset of 3000 frames of the original raw 256×256 NIR images for this part of our study. These images were augmented using TensorFlow and the Keras API, which allowed us to apply random transformations to the dataset. The transformations included random rotation followed by a horizontal flipping, and up to 2% rescaling.

3.15. Training and Testing Dataset

The final training and testing dataset for the machine learning models was determined by the outcome of the congenic pair selection and feature selection steps. This dataset included all congenic groups except for Dll4+ and Dll4–, as well as the selected CG#+ and CG#– groups in the previous step. This step was conducted separately for the original dataset and augmented dataset. The models were trained using 10-fold cross-validation and a portion of the training data was reserved for testing after hyperparameter optimization, with 25% for the original dataset and 20% for the augmented dataset. The performance of the models was assessed using a confusion matrix, receiver operating characteristic curve (ROC), and the area under the curve.

3.16. Statistical Analysis

Repeated measures models are a powerful tool in statistical analysis that allow researchers to study the effects of different factors on a given outcome while accounting for the inherent dependence of multiple measurements taken on the same subject. In this study, a mixed effects model with appropriate time varying covariates was used to analyze the average fluorescence intensity of indocyanine green (ICG) in the tumor with multiple measurements per subject, with the subject number serving as the repeated measure indicator and the rat strain serving as a covariate. This allows for flexible time-based modeling when using multiple measures, likely dependent from the same animal [63,64]. Customized scripts in MATLAB were used to generate the fitted coefficients, covariance parameters, design matrix, error degrees of freedom, and between- and within-subjects factor names for the repeated measures model. The output was then analyzed with a multiple comparison of the estimated marginal means based on the variable strain, using the Tukey-Kramer test statistic [65]. This allowed estimation of multiplicity-adjusted *p*-values for the post hoc comparisons, which indicate whether the groups significantly differed with respect to strain. The data were then visualized as a *p*-value matrix, providing a clear illustration of the significant differences between groups.

3.17. Data Availability

This study employed the established consomic rat models SS and SS.BN3 as well as our congenic strains CG1 to CG8. The publicly accessible and NIH-supported Rat Genome Database (rgd.mcw.edu) catalogs has tools to explore the genotype and pheno-type information for the SS (Dll4+) and SS.BN3 (Dll4–) and congenic strains under strain records: Dll4+ (RGDID:61499), Dll4– (RGDID:1358154), CG1 (RGDID:155782881), CG2 (RG-DID:155782883), CG3 (RGDID:155782884), CG4 (RGDID:155791428), CG5 (RGDID:155791426), CG6 (RGDID:155791430), CG7 (RGDID:155791429), and CG8 (RGDID:155791427).

4. Results and Discussion

4.1. Dynamic Contrast-Enhanced NIR Fluorescence Imaging and Tumor Detection

Dynamic contrast-enhanced NIR fluorescence imaging has been widely used for tumor detection in various studies [66–68]. The use of NIR imaging allows for the visualization of internal organs and tissues without the need for invasive procedures, which can be particularly useful in detecting tumors due to their vascular heterogeneity compared to surrounding healthy tissues.

In previous studies, the use of principal component analysis (PCA) on the time domain of dynamic fluorescence images was utilized to extract spatial patterns of internal organs linked to statistically similar kinetic behaviors, such as liver, kidneys, lungs, and various tumors [7,56]. However, this technique required manual inspection and selection of proper principal components, which was time consuming and prone to human error and bias.

In order to overcome the limitations present in the current dataset, we implemented a modified method that utilizes near-infrared imaging and principal component analysis to detect tumors with high accuracy and without the need for manual correction (Figures 2 and S1). The use of principal component analysis in this context not only allows for dimensionality reduction and noise removal but also enhances the robustness and efficiency of the method.

Our study also implemented a novel method of ranking PCA components based on the 2D cross-correlation of a reference image containing the tumor. This added to the simplicity and computational efficiency of the framework. However, it should be noted that this method may not be effective for detecting tumors with random locations. On the other hand, it could be useful for detecting tumors or tissues of interest with high localization, such as the lungs, liver, and kidney, and lesions in breast tissue. Overall, our method shows potential for improving the accuracy and efficiency of tumor detection using NIR imaging and PCA (Figure 4a). However, further experimentation is needed to expand the framework to a general tumor detection algorithm.



Figure 4. (a) Color-coded and merged principal component images were used to anatomically segment tumors from a single view/projection for rat strains with high (red) or low (black) levels of Dll4 expression. The top 3 ranked PCs for detecting tumors are merged as red, green, and blue channels of a true color (RGB) image. Tumors are outlined in yellow. (**b**–**d**) Mean intensity of ICG biodistribution fluorescence kinetic data over 140 s for (**b**) parental (Dll4+) and consomic (Dll4–) rats, (**c**) all Dll4–high and Dll4–low rats, and (**d**) all strains, plotted and normalized to a range of 0 to 1. Shaded regions show s.e.m.

4.2. Dll4 and Its Effect on the NIR Time Series

The analysis of the average fluorescence intensity of indocyanine green (ICG) in the tumor tissue of Dll4+ and Dll4– rats bearing triple-negative breast cancer (TNBC) tumors revealed that ICG uptake occurred more rapidly in Dll4– tissues and was retained for longer periods of time compared to Dll4+ hosts (Figure 4b). This indicates systemic differences in vascular function between the two rat strains. Our previous histological data showed that Dll4+ tumors have a higher vascular density and tortuosity, indicating a genetic microenvironment that promotes nonproductive angiogenesis [38]. This is further supported by the slower ICG wash-out observed in the Dll4+ tumors. These findings provide insight into the effects of host genetics on tumor angiogenesis and suggest potential therapeutic targets for TNBC.

In order to further investigate the role of Dll4 in vascular function in tumors, we divided chromosome 3 into regions with and without the Dll4 gene in congenic rat strains (Figure 1b) and then examined the ICG fluorescence intensity of tumors in Dll4-high and Dll4-low rats (Video S2) (Figure 4c). Our findings reveal significant systemic differences in vascular function between tumors in Dll4+ and Dll4- rats (parental strains), indicating the critical role of the Dll4 gene in tumor angiogenic response [38]. However, analysis of the ICG fluorescence intensity of tumors for individual strains (Figure 4d) reveals more complex behavior than the obvious differences in wash-in and wash-out patterns observed between Dll4+ and Dll4-. This supports the need for further investigation into the impact of Dll4 on NIR time series signatures and the potential use of Dll4-directed therapies for cancer treatment.

It is worth noting that although there are significant differences in Dll4-low vs. Dll4high rat strains (when all the strains of Dll4 expression levels are combined), they are inconsistent with the observations made in Dll4+ and Dll4- rats. These results have significant implications for developing novel therapies that target Dll4 and other host TME modifiers involved in angiogenesis, as they demonstrate the critical role of these genes in

Additionally, our research further highlights the capricious nature of the NIR signal, which is influenced by various heritable tumor microenvironments across different groups, as shown in Figure 4b,c. We aim to illustrate and categorize the impact of the Dll4 expression level on the NIR signal through this erratic behavior.

We used a repeated measures model to analyze the average fluorescence intensity of ICG in the tumor over time, with the rat strain serving as a covariate. Figure 5a,b show the estimated response covariances matrix, which is the covariance of the repeated measures. The higher values in this matrix indicate the time points at which groups experience the greatest differences. By projecting the diagonal of the covariance matrix onto the time axis (Figure 5c), we were able to visualize the amount of difference between groups over time.



tumor vascular function and angiogenic response.

Figure 5. (**a**,**b**) The covariance parameter of a mixed-effects model for 1400 timesteps with appropriate time-varying covariates, used to analyze the average fluorescence intensity of ICG in the tumor with multiple measurements per subject, with the subject number serving as the repeated measure indicator and the rat strain serving as a covariate. (**c**) Projection of the diagonal of the covariance parameter on the time axis. (**d**) Multiplicity-adjusted *p*-value matrix showing significant differences between groups based on multiple comparison of estimated marginal means using the Tukey–Kramer test statistic. (**e**) Sankey diagram showing differences in strains with high and low levels of Dll4 expression based on the *A*_{score} (mean of accuracy, sensitivity, and specificity in the primary classification of each pair); 0.65 was subtracted from all the *A*_{score} values for better visualizations of differences.

This projection, when compared to the average fluorescence intensity of ICG in the tumor (Figure 4b,d), showed the strongest differences between groups at the points where the NIR signal regions from half of its peak value to the peak value and at the tail of the curve, which are measures of the temporal inhomogeneity of the initial uptake and the decay of ICG fluorescence, were found to be particularly useful in discriminating between groups with different levels of Dll4 expression. This projection of the diagonal of the estimated response covariances matrix on the time curve can be used in feature design to focus on regions with the maximum amount of useful information for discriminating between groups and, subsequently, between classes with different levels of Dll4 expression. This could potentially improve the accuracy of tumor classification and ultimately improve therapy outcomes.

Our repeated measures model, which included responses as measurements and strains as predictor variables, allowed us to conduct multiple comparisons of estimated marginal means between groups. The resulting *p*-value matrix (Figure 5d) revealed significant differences in estimated marginal means between the Dll4+ and Dll4– groups, with a *p*-value of 4.71×10^{-7} . In addition, we observed significant differences between Dll4+ and CG3, CG4, and CG8, with *p*-values of 1.67×10^{-5} , 7.03×10^{-7} , and 2.18×10^{-3} , respectively.

For each group pair with high and low Dll4 expression levels, the separation score was calculated. First each of Dll4+ | CG#-, CG#+ | Dll4-, and CG#+ | CG#- went through our classification algorithm with 10-fold cross-validation using Nearest Neighborhood, Linear SVM, RBF SVM, Decision Tree, Naive Bayes, and Logistic Regression models. The highest average classification metrics (Accuracy + Specificity + Sensitivity)/3 for Dll4+ | CG#-, CG#+ | Dll4- and CG#+ | CG#-) was used to calculate the separation score (Score Dll4+ | CG#- + Score CG#+ | Dll4- + 2 × Score CG#+ | CG#-)/4.

Furthermore, our analysis showed significant differences between Dll4– and CG5 and CG6, with *p*-values of 8.70×10^{-3} and 2.58×10^{-4} , respectively. This supports the hypothesis that Dll4 expression levels can act as a heritable TME modifier on NIR time series intensity. However, the smallest *p*-value between Dll4+ and Dll4– suggests that there are other factors on chromosome 3, in addition to Dll4, that contribute to the observed differences in the NIR time series signature between these groups.

In contrast, no significant differences were found between Dll4– and the congenic strains with low levels of Dll4 expression (CG2, CG3, CG4, CG7, and CG8). This further supports the notion that Dll4 plays a crucial role in determining tumor vascular function and NIR time series intensity.

Among the congenic groups, the most significant differences were observed between CG5, CG6, and CG3, CG4 from the Dll4-high and Dll4-low groups, respectively. Notably, the differences were most significant between CG4 and CG6, with a *p*-value of 0.0003. This suggests that very narrow regions of differences on chromosome 3 between these two groups, one containing Dll4 and the other lacking it, have a significant effect on the NIR time series signature.

4.3. Primary Classification and Congenic Dissimilarity

The relationship between Dll4 expression and classification performance was analyzed using a total of 12 congenic pairs with n > 5 based on their levels of Dll4 expression (Dll4+|Dll4-,Dll4+|CG3,Dll4+|CG4,Dll4-|CG1,Dll4-|CG5,Dll4-|CG6,CG1|CG3,CG1|CG4,CG3|CG5,CG3|CG6,CG4|CG5, and CG4|CG6). The pairs were then subjected to a primary classification algorithm and their mean performance metrics, the A_{score} , were calculated and reported in Table 1. The congenic pairs with low levels of Dll4 expression showed a mean A_{score} of 0.91 +/- 0.01, indicating a high level of classification performance when compared to the Dll4+ parental strain. In contrast, the congenic pairs with high levels of Dll4 expression showed a mean A_{score} of 0.8 +/- 0.05 when classified against the Dll4- consomic strain. Among the congenic pairs, the CG5|CG4 pair demonstrated the highest A_{score} of 0.8, followed by the CG6|CG4 and CG6|CG3 pairs with A_{score} values of 0.78 and 0.77, respectively. The results of the A_{score} calculation are visualized in Figure 5e through a Sankey diagram.

To account for potential differences between the congenic pairs and the parental pairs, the S_{score} was calculated. The CG5 | CG4, CG6 | CG3, and CG6 | CG4 pairs showed the highest S_{score} values of 0.84, 0.84, and 0.85, respectively, and were selected for the feature selection step. These results align with the multiple comparison of estimated marginal means between groups, indicating that CG5 | CG4, CG6 | CG3, and CG6 | CG4 show the strongest differences in classification performance.

4.4. Feature Selection

RFE is a wrapper method that evaluates the entire classification algorithm and has shown improved classification accuracy and reduced overfitting compared to other feature selection methods [69]. However, RFE can be sensitive to noise and irrelevant features, leading to suboptimal feature subsets and reduced classification performance. Additionally, RFE is computationally intensive, which can pose a challenge for large datasets with a high number of features. Despite these limitations, RFE remains a valuable tool for selecting an optimal subset of features that maximizes classification performance [70,71].

To address these limitations, we performed feature selection in two steps to optimize the selection process and improve the performance of the classifier. First, we used RFE to select only two features out of the 86 available features for congenic pairs Dll4+ | Dll4-, Dll4+ | CG4, CG5 | Dll4-, CG5 | CG4, CG6 | Dll4-, and CG6 | CG4. The CG3 and its combinations (CG6 | CG3, CG5 | CG3, and Dll4+ | CG3) were dropped from the feature selection process as the performances of the classifiers, the A_{score} , using only two features were below 0.70, and lower than the other strains. The congenic pairs CG5 | CG4 and CG6 | CG4 as well as the parental and consomic group Dll4+ | Dll4- went through our feature selection algorithm, and for each pair the two best-performing classification algorithms based on A_{score} and associated feature pair were reported (Table 2).

Creation	Feature	Best Average		Best Accuracy		Best Sensitivity		Best Specificity	
Groups		Alg.	Value(std)	Alg.	Value(std)	Alg.	Value(std)	Alg.	Value(std)
	HIF10_avg, HIF12_avg	DT	0.8666 (0.2309)	DT	1 (0)	DT	0.6 (0.5026)	DT	1 (0)
D114++D114-	HIF6_avg, HIF50_avg	LR	0.7597 (0.1412)	LR	0.8682 (0.1468)	LR	0.6 (0.5026)	DT	0.9391 (0.116)
	HIF300_avg, HIF200_avg	RBF SVM	0.6415 (0.3196)	RBF SVM	0.5914 (0.0983)	L SVM	1 (0)	KNN	0.9166 (0.1147)
DII4++CG3	HIF300_avg, TR_rel	LR	0.6453 (0.1036)	LR	0.6176 (0.1624)	LR	0.85 (0.2665)	KNN	0.9083 (0.1147)
	HIF30_avg, D16_avg	DT	0.9175 (0.0303)	DT	0.9125 (0.0915)	L SVM	0.95 (0.0888)	NB	1 (0)
DII4++CG4	HIF8_avg, D6_avg	RBF SVM	0.88125 (0.1254)	RBF SVM	0.89375 (0.1174)	L SVM	1 (0)	NB	0.85 (0.2016)
CG5 Dll4-	HIF5_avg, HIF16_avg	DT	0.8251 (0.0698)	DT	0.8292 (0.1093)	L SVM	1 (0)	NB	0.845 (0.1952)
	HIF8_avg, HIF25_avg	KNN	0.8044 (0.1505)	KNN	0.8201 (0.0919)	L SVM	1 (0)	NB	0.7925 (0.2014)
CCELCCO	HIF300_avg, HIF400_avg	L SVM	0.5683 (0.4037)	LR	0.5292 (0.1507)	L SVM	1 (0)	KNN	0.85 (0.1613)
CG51CG3	HIF300_avg, HIF50_avg	LR	0.5919 (0.4037)	LR	0.5626 (0.1652)	L SVM	1 (0)	KNN	0.8333 (0.1324)
CG5 CG4	D18_avg, HIF4_rel	KNN	0.7811 (0.0485)	KNN	0.7948 (0)	L SVM	0.8928 (0)	DT	0.7272 (0)
	D18_avg, HIF4_avg	KNN	0.7811 (0.0485)	KNN	0.7948 (0)	L SVM	0.8928 (0)	DT	0.7272 (0)
	D40_avg, D2_avg	KNN	0.8685 (0.0404)	KNN	0.8755 (0.1180)	L SVM	0.95 (0.1574	KNN	0.905 (0.1422)
CG6 DII4-	D40_avg, HIF5_rel	KNN	0.8482 (0.01263)	KNN	0.8440 (0.1386)	RBF SVM	0.9375 (0.1293)	KNN	0.8383 (0.1643)
CG61CG3	HIF300_avg, HIF350_avg	L SVM	0.5729 (0.3995)	KNN	0.5318 (0.1215)	L SVM	1 (0)	NN	0.8583 (0.1733)
	HIF300_avg, HIF180_avg	RBF SVM	0.5890 (0.2399)	RBF SVM	0.5570 (0.1259)	L SVM	1 (0)	KNN	0.85 (0.2222)
00(1001	HIF8_avg, HIF23_avg	KNN	0.7278 (0.1917)	KNN	0.7437 (0.0904)	RBF SVM	1 (0)	NB	0.7285 (0.2853)
CG61CG4	HIF6_avg, HIF50_avg	NB	0.7280 (0.2263)	NB	0.7468 (0.0982)	L SVM	1 (0)	DT	0.6071 (0.1846)

Table 2. Primary validation and feature selection.

For each group pair with high and low Dll4 expression levels, 2 sets of feature pairs and best-performing algorithms listed. For each algorithm, accuracy, sensitivity, specificity metrics, and average metric (Accuracy + Specificity + Sensitivity)/3 shown. Alg.: classifier algorithm, KNN: Nearest Neighborhood, L SVM: Linear SVM, RBF SVM: Radial Basis Function kernel SVM, DT: Decision Tree, NB: Naive Bayes, LR: Logistic Regression models.

The highest A_{score} for the two best-performing models for CG5 | CG4 was 0.78 ± 0.04 compared to CG6 | CG4 with an A_{score} of 0.72 ± 0.19 and 0.72 ± 0.22, resulting in the selection of CG5 | CG4 for final congenic pair selection.

Finally, from each pair of Dll4+ | Dll4-, Dll4+ | CG4, CG5 | Dll4, and CG4 | CG5, four of the best performing features regardless of the ML model were chosen and were used as a collection of features for the final feature selection (Table 2). A combination of parental and consomic groups and the final selected congenic pair (Dll4+, Dll4-, CG5, and CG4)

was used to select the final feature pair out of the 16 selected features, resulting in the selection of HIF5_avg and HIF50_avg as the best-performing features.

4.5. Performance of the Classification Models Based on the Selected Features

To evaluate the performance of the selected features, we trained datasets consisting of all the remaining congenic groups excluding the Dll4+, Dll4–, CG4, and CG5 (CG1 to CG3 and CG5 to CG8) using 10-fold cross-validation and keeping 25% of the dataset for testing the trained models. This allowed us to assess the generalizability of our model and test it on previously unseen datasets. The results of this step are reported as a confusion matrix, ROC curve, and AUC (Figure 6), as well as general classification metrics (Table 3). The best-performing models based on the selected features were SVM and KNN, with sensitivity and specificity of 1.00 and 0.81 and 1.00 and 0.75, respectively.



Figure 6. Performance of selected features on original dataset. The confusion matrix and ROC plots for training and testing of best-performing models based on HIF5_avg and HIF50_avg features with 10-fold cross-validation. The dataset includes all rat strains except Dll4+, Dll4-, CG4-, and CG5+. An amount of 25% of the dataset was used for holdout testing. Top row: SVM model. Bottom row: Nearest Neighborhood.

In order to further assess the effectiveness of our model, the selected features, and the generated congenic pair, we generated an augmented dataset consisting of all remaining congenic pairs excluding Dll4+, Dll4–, CG4 and CG5 (CG1 to CG3 and CG5 to CG8, with random variations in rotation, horizontal flip, and limited scaling (up to $\pm 2\%$) to increase the diversity of the dataset. This resulted in a total of 606 data points. The performance of the models was evaluated using 10-fold cross-validation and a 20% hold out. The results of this step were reported as a confusion matrix, ROC curve, AUC (Figure 7), and overall classification metrics (Table 4). The best-performing models based on the selected features were SVM and KNN, with sensitivity and specificity of 0.97 and 0.91, and 0.97 and 0.92, respectively. These results align closely with the performance of the models over the original dataset, indicating the generalizability of our framework.

	SVM		KN	N
Measure	Training	Testing	Training	Testing
Sensitivity	0.9310	1.0000	0.8108	1.0000
Specificity	0.8182	0.8182	0.9286	0.7500
Precision	0.8710	0.7500	0.9677	0.6250
Negative Predictive Value	0.9000	1.0000	0.6500	1.0000
False-Positive Rate	0.1818	0.1818	0.0714	0.2500
False Discovery Rate	0.1290	0.2500	0.0323	0.3750
False-Negative Rate	0.0690	0.0000	0.1892	0.0000
Accuracy	0.8824	0.8824	0.8431	0.8235
F1 Score	0.9000	0.8571	0.8824	0.7692
Matthews Correlation Coefficient	0.7600	0.7833	0.6758	0.6847

 Table 3. Performance of selected features.

The training and testing metrics of the best-performing models based on the selected features HIF5_avg and HIF50_avg with 10-fold cross-validation are shown. The training dataset includes features extracted from the original NIR time series data of all congenic groups except CG5+ and CG4-, and 25% of the dataset was used for holdout testing of the models. SVM: Support Vector Machine, KNN: Nearest Neighborhood.



Figure 7. Performance of selected features on augmented dataset. The confusion matrix and ROC plots for training and testing of best-performing models based on HIF5_avg and HIF50_avg features with 10-fold cross-validation. The dataset includes all rat strains except Dll4+, Dll4-, CG4-, and CG5+. 25% of the dataset was used for holdout testing. Top row: SVM model. Bottom row: Nearest Neighborhood.

 Table 4. Performance of selected features on augmented dataset.

	SVM		KKN	
Measure	Training	Testing	Training	Testing
Sensitivity	0.9253	0.9756	0.9702	0.9762
Specificity	0.9132	0.9125	0.9211	0.9241
Precision	0.8564	0.8511	0.867	0.8723
Negative Predictive Value	0.9562	0.9865	0.9832	0.9865
False Positive Rate	0.0868	0.0875	0.0789	0.0759
False Discovery Rate	0.1436	0.1489	0.133	0.1277
False Negative Rate	0.0747	0.0244	0.0298	0.0238
Accuracy	0.9175	0.9339	0.9381	0.9421
F1 Score	0.8895	0.9091	0.9157	0.9213
Matthews Correlation Coefficient	0.8254	0.8625	0.8705	0.8793

The training and testing metrics of the best-performing models based on the selected features HIF5_avg and HIF50_avg with 10-fold cross-validation are shown. The training dataset includes features extracted from the original NIR time series data of all congenic groups except CG5+ and CG4-, and 25% of the dataset was used for holdout testing of the models. SVM: Support Vector Machine, KNN: Nearest Neighborhood.

It is noteworthy that of the 16 most contributing features used to select the final feature pair, 12 were the newly proposed HIF features, and the other 4 were DS features, which we previously reported [48]. Additionally, the HIF features were amongst the best features for identifying genetic TME modifiers.

The relationship between covariance of the repeated measures and optimal feature design in machine learning classification algorithms is an essential factor in developing effective classification algorithms. Combined with our recent report [47], our analysis found that the DS and HIF features, which are generated in regions where the NIR signal varies from half of its peak value to the peak value and at the tail of the curve, were particularly effective in discriminating between benign/malignant tumors (DS features) and groups with different levels of Dll4 expression (HIF features). Furthermore, the projection of the covariance matrix onto the time axis revealed similar regions, indicating a relationship between this projection and optimal feature design. These findings have significant implications for feature design in machine learning classification algorithms. By focusing on the regions with the greatest amount of useful information for discrimination, we can design features specifically to capture these differences and improve the accuracy of tumor classification. This can ultimately lead to better therapy outcomes for patients. It is worth noting that this relationship between the covariance of the repeated measures and optimal feature design is not limited to HIF features and the specific context of our analysis. In general, considering the covariance of repeated measures can provide valuable information for identifying key regions and designing effective features for machine learning classification algorithms.

5. Conclusions

Dynamic vascular imaging techniques such as DCE-MRI and perfusion CT are used to extract multiple vascular parameters and have been used in clinical trials of anti-angiogenic drugs. However, these techniques have limitations, such as low temporal resolution and the need for specialized hardware and software [41,42]. To overcome these limitations, dynamic near-infrared (NIR) fluorescence imaging can serve as an effective alternative for characterizing germline-dependent vascular phenotypes. It can be combined with other modalities, such as in a paired-agent or multimodal MRI and fluorescence tomography approaches for noninvasive quantification of response to anti-angiogenesis therapy and classifying in vivo vascular phenotypes [7,43,44,67]. Furthermore, DE-NIR imaging, as a potential alternative for characterizing germline-dependent vascular germline-dependent vascular phenotypes in preclinical models, can be extended to clinical modalities upon validation with cross-sectional dynamic contrast enhanced imaging.

The present study proposes that by combining DE-NIR imaging and machine learning algorithms with consomic xenograft models with human tumors, the role of inherited notch protein Dll4 (rat variant of delta like canonical ligand 4) expression specifically in the host vascular microenvironment can be studied. Specifically, in the context of breast cancer, where different genetic subtypes can impact treatment outcomes, identifying patients with high or low DLL4 (human variant of delta like canonical ligand 4) expression levels through noninvasive imaging could assist in selecting personalized treatment options. Nonetheless, the study authors acknowledge notable differences between the rat model utilized in the study and the human system, which could affect the generalization of the findings to human cases, as in human tumors DLL4 expression maybe be both on tumor cells and host vasculature, whereas in our CXM model, we focused specifically on the inherited variation in rat-derived host vasculature Dll4 expression in human xenograft tumors. Such differences may result in amplification or suppression of vascular phenotype responses if both tumor cells and the host microenvironment express high levels or contrasting levels of DLL4. However, even in that case, dynamic imaging will be useful in identifying patients likely to respond better to DLL4 targeted therapies.

Future studies will be necessary to validate this study's findings, to assess the reliability and validity of the developed imaging and machine learning algorithms in a large and diverse patient population to determine if contrast agent kinetic profiles observed in human DCE-MRI or dyna-CT imaging datasets for primary and/or metastatic disease differ in human patients with high vs. low DLL4 expression. In the metastatic setting, where surgery is no longer an option, a machine-learning-enabled analysis of dynamic contrast-enhanced imaging will be valuable to assess the expression levels of DLL4 and guide therapy selection, especially in cases where a biopsy is not taken or if biopsy results are inconclusive [72,73]. Human anti-DLL4 antibodies have been reported for cancer treatment [54,74–76]. In one study, immunotoxin DLL4Nb-PE was developed, potentially as a cell cytotoxic agent and angiogenesis maturation inhibitor [76]. Another study successfully developed a bispecific monoclonal antibody that targets both human DLL4 and VEGF and showed efficacy in inhibiting proliferation, migration, and tube formation of human umbilical vein endothelial cells (HUVEC) [74]. In a phase 1a trial, navicixizumab, a bispecific antibody that inhibits DLL4 and VEGF, was tested in refractory solid tumor patients and showed the potential to inhibit tumor growth [75]. While DLL4 blockade is an attractive therapy, long-term extended use of DLL4 mAbs has demonstrated concerning off-target effects [77,78]. Pharmacokinetic modulation of DLL4 mAbs may reduce off-target effects [77], such as via short-term administration or by focusing on patients where dynamic contrast imaging indicates a high DLL4 vascular phenotype. As we have shown in prior work, high vascular DLL4 expressing tumors may also be susceptible to DLL4 targeted nanomedicine [38] or as combination therapy with anti-DLL4 monoclonal antibodies with nanomedicine drugs such as nab-paclitaxel (Abraxane) or Liposomal Doxorubicin (DoxilTM).

The use of noninvasive DE-NIR imaging to detect heritable TME modifiers is significant for several reasons. First, this method allows for the identification of potential modifiers without the need for invasive procedures, reducing the potential for discomfort and complications for patients. Second, the use of machine learning and DE-NIR imaging to develop a predictive model for cancer nanomedicine therapy can support effective decision making in the treatment process. While data processing and preparation and algorithm training can be complex, the resulting algorithms are simple and allow for the prediction of heterogeneity in a single step using ROI brightness measurements. Interestingly, traditional features such as time-to-peak and upslope do not appear in our selection of the most discriminative features. However, two novel features derived from *H1F* (HIF5_avg and HIF50_avg), which is a measure of the temporal inhomogeneity of both the initial uptake and decay of ICG fluorescence, were identified.

It is important to note that the training and testing sets used in this study are minimal, and therefore the high accuracy rates obtained should be interpreted with caution. Further research with larger datasets will be necessary to assess the reliability and validity of these findings with confidence.

We have reported novel dynamic enhanced near-infrared (NIR) fluorescence imaging and machine learning algorithms to noninvasively assess Dll4 expression levels in tumors. Our results showed that observation of subtle differences in vasculature structure and perfusion patterns characterized by ICG time kinetics could be used to differentiate between inherited tumor vascular microenvironment differences, such as Dll4 expression levels. Additionally, our analysis demonstrated the importance of considering the covariance of the repeated measures in the design of features for machine learning classification algorithms. By utilizing this information, we can improve the accuracy of tumor classification and ultimately improve therapy outcomes for patients.

To summarize, based on our recent study, we investigated the impact of genetically heterogeneous notch-Dll4 inheritance on the contrast agent uptake and clearance in triplenegative breast cancer xenografts. The differences in Dll4 inheritance have been shown to impact nanomedicine biodistribution, pharmacokinetics, and therapy response in our prior work. Thus, our results indicated that imaging can be potentially employed for selecting patients for Dll4-directed therapies by identify host microenvironments with highor low-expressing Dll4 inheritance. This further suggests that the success of nanomedicine might depend on hereditary tumor microenvironment genes, regardless of tumor type. Additionally, host genes such as Dll4 can affect individual differences in NP uptake and response to NP-mediated therapies, providing the potential for more effective personalized Dll4 targeted nanomedicine for therapy-resistant hosts. Further studies are needed to validate these findings and explore the potential clinical applications of this approach.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/cancers15051460/s1, Figure S1: Principal component analysis; Video S1: RespMotionCorretion; Video S2: ICG_NIR_All_Strains.

Author Contributions: Conceptualization, S.S. and A.J.; Data curation, S.S. and J.J.; Formal analysis, S.S. and A.J.; Funding acquisition, A.J.; Investigation, S.S.; Methodology, S.S., A.B., A.G., M.F. and A.J.; Project administration, A.J.; Resources, A.J.; Software, S.S., M.Z. and J.E.; Supervision, S.Z. and A.J.; Validation, S.S.; Visualization, S.S.; Writing—original draft, S.S.; Writing—review and editing, J.E., A.B. and A.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by funding from NIH/NCI 2R01CA193343 to the MCW Cancer Center.

Institutional Review Board Statement: All methods have been carried out in accordance with relevant guidelines and regulations. Approved protocols by the Medical College of Wisconsin Institutional Biosafety Committee (IBC) and Institutional Animal Care and Use Committee (IACUC) were followed. All live animal experiments are reported per the ARRIVE guidelines' recommendations.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study employed the established consomic rat models SS and SS.BN3 as well as our congenic strains CG1 to CG8. The publicly accessible and NIH-supported Rat Genome Database (rgd.mcw.edu) catalogs has tools to explore the genotype and phenotype information for the SS (Dll4+) and SS.BN3 (Dll4–) and congenic strains under strain records Dll4+ (RGDID:61499), Dll4– (RGDID:1358154), CG1 (RGDID:155782881), CG2 (RGDID:155782883), CG3 (RGDID:155782884), CG4 (RGDID:155791428), CG5 (RGDID:155791426), CG6 (RGDID:155791430), CG7 (RGDID:155791429), and CG8 (RGDID:155791427).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Pasha, N.; Turner, N.C. Understanding and overcoming tumor heterogeneity in metastatic breast cancer treatment. *Nat. Cancer* 2021, 2, 680–692. [CrossRef]
- Marisa, L.; de Reyniès, A.; Duval, A.; Selves, J.; Gaub, M.P.; Vescovo, L.; Etienne-Grimaldi, M.-C.; Schiappa, R.; Guenot, D.; Ayadi, M.; et al. Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value. *PLoS Med.* 2013, *10*, e1001453. [CrossRef]
- Sohn, B.H.; Hwang, J.-E.; Jang, H.-J.; Lee, H.-S.; Oh, S.C.; Shim, J.-J.; Lee, K.-W.; Kim, E.H.; Yim, S.Y.; Lee, S.H.; et al. Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. *Clin. Cancer Res.* 2017, 23, 4441–4449. [CrossRef]
- Lehmann, B.D.; Pietenpol, J.A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. J. Pathol. 2014, 232, 142–150. [CrossRef]
- Belli, C.; Trapani, D.; Viale, G.; D'Amico, P.; Duso, B.A.; Della Vigna, P.; Orsi, F.; Curigliano, G. Targeting the microenvironment in solid tumors. *Cancer Treat. Rev.* 2018, 65, 22–32. [CrossRef]
- Bernard, J.J.; Wellberg, E.A. The Tumor Promotional Role of Adipocytes in the Breast Cancer Microenvironment and Macroenvironment. Am. J. Pathol. 2021, 191, 1342–1352. [CrossRef]
- Jagtap, J.; Sharma, G.; Parchur, A.K.; Gogineni, V.; Bergom, C.; White, S.; Flister, M.J.; Joshi, A. Methods for detecting host genetic modifiers of tumor vascular function using dynamic near-infrared fluorescence imaging. *Biomed. Opt. Express* 2018, *9*, 543–556. [CrossRef]
- 8. Baghban, R.; Roshangar, L.; Jahanban-Esfahlan, R.; Seidi, K.; Ebrahimi-Kalan, A.; Jaymand, M.; Kolahian, S.; Javaheri, T.; Zare, P. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun. Signal.* **2020**, *18*, 59. [CrossRef]
- 9. Economopoulou, P.; Kotsantis, I.; Psyrri, A. Tumor Microenvironment and Immunotherapy Response in Head and Neck Cancer. *Cancers* 2020, *12*, 3377. [CrossRef]
- Hinshaw, D.C.; Shevde, L.A. The tumor microenvironment innately modulates cancer progression. *Cancer Res.* 2019, 79, 4557–4567. [CrossRef]
- 11. Pitt, J.M.; Marabelle, A.; Eggermont, A.; Soria, J.-C.; Kroemer, G.; Zitvogel, L. Targeting the tumor microenvironment: Removing obstruction to anticancer immune responses and immunotherapy. *Ann. Oncol.* **2016**, *27*, 1482–1492. [CrossRef]

- 12. Ungefroren, H.; Sebens, S.; Seidl, D.; Lehnert, H.; Hass, R. Interaction of tumor cells with the microenvironment. *Cell Commun. Signal.* **2011**, *9*, 18. [CrossRef]
- Stamatelos, S.K.; Bhargava, A.; Kim, E.; Popel, A.S.; Pathak, A.P. Tumor Ensemble-Based Modeling and Visualization of Emergent Angiogenic Heterogeneity in Breast Cancer. Sci. Rep. 2019, 9, 5276. [CrossRef]
- Stamatelos, S.; Kim, E.; Pathak, A.; Popel, A. Image-based Characterization of Functional and Structural Heterogeneity of Tumor Xenografts using Blood Flow modeling, Oxygenation Modeling and Multivariate Analysis. FASEB J. 2015, 29, 787.11. [CrossRef]
- 15. Jain, R.K. Normalizing Tumor Microenvironment to Treat Cancer: Bench to Bedside to Biomarkers. J. Clin. Oncol. 2013, 31, 2205–2218. [CrossRef]
- 16. Noman, M.Z.; Hasmim, M.; Messai, Y.; Terry, S.; Kieda, C.; Janji, B.; Chouaib, S. Hypoxia: A key player in antitumor immune response. A Review in the Theme: Cellular Responses to Hypoxia. *Am. J. Physiol. Cell Physiol.* **2015**, *309*, C569–C579. [CrossRef]
- 17. Muz, B.; de la Puente, P.; Azab, F.; Azab, A.K. The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Hypoxia* **2015**, *3*, 83–92. [CrossRef]
- Willett, C.G.; Duda, D.G.; di Tomaso, E.; Boucher, Y.; Ancukiewicz, M.; Sahani, D.V.; Lahdenranta, J.; Chung, D.C.; Fischman, A.J.; Lauwers, G.Y.; et al. Efficacy, Safety, and Biomarkers of Neoadjuvant Bevacizumab, Radiation Therapy, and Fluorouracil in Rectal Cancer: A Multidisciplinary Phase II Study. J. Clin. Oncol. 2009, 27, 3020–3026. [CrossRef]
- Raut, C.P.; Boucher, Y.; Duda, D.G.; Morgan, J.A.; Quek, R.; Ancukiewicz, M.; Lahdenranta, J.; Eder, J.P.; Demetri, G.D.; Jain, R.K. Effects of Sorafenib on Intra-Tumoral Interstitial Fluid Pressure and Circulating Biomarkers in Patients with Refractory Sarcomas (NCI Protocol 6948). PLoS ONE 2012, 7, e26331. [CrossRef]
- Eichler, A.F.; Chung, E.; Kodack, D.P.; Loeffler, J.S.; Fukumura, D.; Jain, R.K. The biology of brain metastases—Translation to new therapies. *Nat. Rev. Clin. Oncol.* 2011, *8*, 344–356. [CrossRef]
- 21. Hobbs, S.K.; Monsky, W.L.; Yuan, F.; Roberts, W.G.; Griffith, L.; Torchilin, V.P.; Jain, R.K. Regulation of transport pathways in tumor vessels: Role of tumor type and microenvironment. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4607–4612. [CrossRef]
- 22. Lee, I.; Demhartner, T.J.; Boucher, Y.; Jain, R.K.; Intaglietta, M. Effect of Hemodilution and Resuscitation on Tumor Interstitial Fluid Pressure, Blood Flow, and Oxygenation. *Microvasc. Res.* **1994**, *48*, 1–12. [CrossRef]
- Vakoc, B.J.; Lanning, R.M.; A Tyrrell, J.; Padera, T.; A Bartlett, L.; Stylianopoulos, T.; Munn, L.L.; Tearney, G.J.; Fukumura, D.; Jain, R.K.; et al. Three-dimensional microscopy of the tumor microenvironment in vivo using optical frequency domain imaging. *Nat. Med.* 2009, 15, 1219–1223. [CrossRef]
- Lugano, R.; Ramachandran, M.; Dimberg, A. Tumor angiogenesis: Causes, consequences, challenges and opportunities. *Cell. Mol. Life Sci.* 2020, 77, 1745–1770. [CrossRef]
- Kazerounian, S.; Lawler, J. Integration of pro- and anti-angiogenic signals by endothelial cells. J. Cell Commun. Signal. 2017, 12, 171–179. [CrossRef]
- 26. Brzozowa-Zasada, M. The role of Notch ligand, Delta-like ligand 4 (DLL4), in cancer angiogenesis—Implications for therapy. *Eur. Surg.-Acta Chir. Austriaca* 2021, 53, 274–280. [CrossRef]
- Canavese, M.; Altruda, F.; Ruzicka, T.; Schauber, J. Vascular endothelial growth factor (VEGF) in the pathogenesis of psoriasis— A possible target for novel therapies? J. Dermatol. Sci. 2010, 58, 171–176. [CrossRef]
- Akil, A.; Gutiérrez-García, A.K.; Guenter, R.; Rose, J.B.; Beck, A.W.; Chen, H.; Ren, B. Notch Signaling in Vascular Endothelial Cells, Angiogenesis, and Tumor Progression: An Update and Prospective. *Front. Cell Dev. Biol.* 2021, 9, 642352. [CrossRef]
- Hellström, M.; Phng, L.-K.; Hofmann, J.J.; Wallgard, E.; Coultas, L.; Lindblom, P.; Alva, J.; Nilsson, A.-K.; Karlsson, L.; Gaiano, N.; et al. Dll4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature* 2007, 445, 776–780. [CrossRef]
- 30. Noguera-Troise, I.; Daly, C.; Papadopoulos, N.J.; Coetzee, S.; Boland, P.; Gale, N.W.; Lin, H.C.; Yancopoulos, G.D.; Thurston, G. Blockade of Dll4 inhibits tumour growth by promoting non-productive angiogenesis. *Nature* **2006**, 444, 1032–1037. [CrossRef]
- 31. Ding, Z.; Zu, S.; Gu, J. Evaluating the molecule-based prediction of clinical drug responses in cancer. *Bioinformatics* **2016**, *32*, 2891–2895. [CrossRef] [PubMed]
- Zohny, S.F.; Zamzami, M.A.; Al-Malki, A.L.; Trabulsi, N.H. Highly Expressed DLL4 and JAG1: Their Role in Incidence of Breast Cancer Metastasis. Arch. Med. Res. 2020, 51, 145–152. [CrossRef]
- A Charles, N.; Holland, E.C. The perivascular niche microenvironment in brain tumor progression. *Cell Cycle* 2010, *9*, 3084–3093. [CrossRef] [PubMed]
- Hyatt, D.C.; Ceresa, B.P. Cellular localization of the activated EGFR determines its effect on cell growth in MDA-MB-468 cells. Exp. Cell Res. 2008, 314, 3415–3425. [CrossRef] [PubMed]
- Scehnet, J.S.; Jiang, W.; Kumar, S.R.; Krasnoperov, V.; Trindade, A.; Benedito, R.; Djokovic, D.; Borges, C.; Ley, E.J.; Duarte, A.; et al. Inhibition of Dll4-mediated signaling induces proliferation of immature vessels and results in poor tissue perfusion. *Blood* 2007, 109, 4753–4760. [CrossRef] [PubMed]
- Hu, W.; Lu, C.; Dong, H.H.; Huang, J.; Shen, D.-Y.; Stone, R.L.; Nick, A.M.; Shahzad, M.M.; Mora, E.; Jennings, N.B.; et al. Biological Roles of the Delta Family Notch Ligand Dll4 in Tumor and Endothelial Cells in Ovarian Cancer. *Cancer Res* 2011, 71, 6030–6039. [CrossRef]
- Comunanza, V.; Bussolino, F. Therapy for Cancer: Strategy of Combining Anti-Angiogenic and Target Therapies. *Front. Cell Dev. Biol.* 2017, 5, 101. [CrossRef] [PubMed]

- Sharma, G.; Jagtap, J.; Parchur, A.K.; Gogineni, V.R.; Ran, S.; Bergom, C.; White, S.B.; Flister, M.J.; Joshi, A. Heritable modifiers of the tumor microenvironment influence nanoparticle uptake, distribution and response to photothermal therapy. *Theranostics* 2020, 10, 5368–5383. [CrossRef] [PubMed]
- 39. Flister, M.J.; Tsaih, S.-W.; Stoddard, A.; Plasterer, C.; Jagtap, J.; Parchur, A.K.; Sharma, G.; Prisco, A.R.; Lemke, A.; Murphy, D.; et al. Host genetic modifiers of nonproductive angiogenesis inhibit breast cancer. *Breast Cancer Res. Treat.* 2017, *165*, 53–64. [CrossRef]
- 40. Flister, M.J.; Endres, B.T.; Rudemiller, N.; Sarkis, A.B.; Santarriaga, S.; Roy, I.; Lemke, A.; Geurts, A.M.; Moreno, C.; Ran, S.; et al. CXM: A New Tool for Mapping Breast Cancer Risk in the Tumor Microenvironment. *Cancer Res* 2014, 74, 6419–6429. [CrossRef]
- 41. Goh, V.; Padhani, A.; Rasheed, S. Functional imaging of colorectal cancer angiogenesis. *Lancet Oncol.* 2007, *8*, 245–255. [CrossRef] [PubMed]
- 42. Choi, M.; Choi, K.; Ryu, S.-W.; Lee, J.; Choi, C. Dynamic fluorescence imaging for multiparametric measurement of tumor vasculature. J. Biomed. Opt. 2011, 16, 046008. [CrossRef] [PubMed]
- 43. Zhu, B.; Godavarty, A. Near-Infrared Fluorescence-Enhanced Optical Tomography. *BioMed Res. Int.* 2016, 2016, 5040814. [CrossRef]
- Meng, B.; Strawbridge, R.R.; Tichauer, K.M.; Samkoe, K.S.; Davis, S.C. Estimating paired-agent uptake in altered tumor vasculature using MRI-coupled fluorescence tomography. *Proc. SPIE Int. Soc. Opt. Eng.* 2020, 11216, 112160U. [CrossRef] [PubMed]
- Chan, J.W.; Hohenstein, N.; Carpenter, C.; Pattison, A.J.; Morin, O.; Valdes, G.; Sanchez, C.T.; Perkins, J.; Solberg, T.D.; Yom, S.S. Artificial Intelligence-Guided Prediction of Dental Doses Before Planning of Radiation Therapy for Oropharyngeal Cancer: Technical Development and Initial Feasibility of Implementation. *Adv. Radiat. Oncol.* 2022, *7*, 100886. [CrossRef] [PubMed]
- Brodin, P.; Shukla, P.; Hauze, M.; Shulte, L.; Carpenter, C.; Kumar, R.; Bodner, W.; Kalnicki, S.; Garg, M.; Tomé, W. PD-0322 Artificial intelligence organ-at-risk dose prediction for high-risk prostate cancer IMRT. *Radiother. Oncol.* 2022, 170, S284–S285. [CrossRef]
- Zhuk, S.; Epperlein, J.P.; Nair, R.; Thirupati, S.; Aonghusa, P.M.; Cahill, R.; O'Shea, D. Perfusion Quantification from Endoscopic Videos: Learning to Read Tumor Signatures. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). *arXiv* 2006. [CrossRef]
- Epperlein, J.P.; Zayats, M.; Tirupathi, S.; Zhuk, S.; Tchrakian, T.; Mac Aonghusa, P.; O'Shea, D.F.; Hardy, N.P.; Dalli, J.; Cahill, R.A. Practical Perfusion Quantification in Multispectral Endoscopic Video: Using the Minutes after ICG Administration to Assess Tissue Pathology. AMIA Annu. Symp. Proc. 2021, 2021, 428.
- 49. Houston, J.P. Near Infrared Optical Lymphography for Cancer Diagnostics. Ph.D. Dissertation, Texas A&M University, Canyon, TX, USA, 2007. Available online: https://oaktrust.library.tamu.edu/handle/1969.1/4807 (accessed on 14 December 2022).
- 50. Flister, M.J.; Hoffman, M.J.; Reddy, P.; Jacob, H.J.; Moreno, C. Congenic Mapping and Sequence Analysis of the *Renin* Locus. *Hypertension* **2013**, *61*, 850–856. [CrossRef]
- 51. Kilkenny, C.; Browne, W.J.; Cuthill, I.C.; Emerson, M.; Altman, D.G. Improving Bioscience Research Reporting: The ARRIVE Guidelines for Reporting Animal Research. *PLoS Biol.* **2010**, *8*, e1000412. [CrossRef]
- Rat Research Model Service Center | Genomic Sciences and Precision Medicine Center | Medical College of Wisconsin. Available online: https://www.mcw.edu/departments/genomic-sciences-and-precision-medicine-center-gspmc/about-us/what-wedo/research/rat-research-model-service-center (accessed on 7 November 2022).
- 53. Flister, M.J.; Joshi, A.; Bergom, C.; Rui, H. Mapping Mammary Tumor Traits in the Rat. *Methods Mol. Biol.* 2019, 2018, 249–267. [CrossRef] [PubMed]
- Miles, K.M.; Seshadri, M.; Ciamporcero, E.; Adelaiye, R.; Gillard, B.; Sotomayor, P.; Attwood, K.; Shen, L.; Conroy, D.; Kuhnert, F.; et al. Dll4 Blockade Potentiates the Anti-Tumor Effects of VEGF Inhibition in Renal Cell Carcinoma Patient-Derived Xenografts. *PLoS ONE* 2014, 9, e112371. [CrossRef] [PubMed]
- 55. Jolliffe, I. Principal Component Analysis for Special Types of Data. 2002. Available online: https://link.springer.com/content/pdf/10.1007/0-387-22440-8_13.pdf (accessed on 9 December 2022).
- Seo, J.; An, Y.; Lee, J.; Ku, T.; Kang, Y.; Ahn, C.; Choi, C. Principal component analysis of dynamic fluorescence images for diagnosis of diabetic vasculopathy. J. Biomed. Opt. 2016, 21, 46003. [CrossRef] [PubMed]
- 57. Benson, R.C.; Kues, H.A. Fluorescence properties of indocyanine green as related to angiography. *Phys. Med. Biol.* **1978**, 23, 159–163. [CrossRef]
- Savitzky, A.; Golay, M.J. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Z. Physiol Chem. 1951, 40, 1832. Available online: https://pubs.acs.org/sharingguidelines (accessed on 23 November 2022). [CrossRef]
- Turner, E. Predictive Variable Selection for Subgroup Identification (Doctoral dissertation, The University of Manchester (United Kingdom)). 2018. Available online: https://www.proquest.com/openview/453e7fde6f6a9539be513ad1523a27f7/1?pq-origsite= gscholar&cbl=2026366 (accessed on 22 October 2022).
- 60. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. Available online: http://scikit-learn.sourceforge.net (accessed on 28 October 2022).
- Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* 2002, 46, 389–422. [CrossRef]

- 62. Cristianini, N.; Shawe-Taylor, J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. 2000. Available online: https://books.google.com/books?hl=en&lr=&id=_PXJn_cxv0AC&oi=fnd&pg=PR9&dq=Cristianini,+N.,+%2 6+Shawe-Taylor,+J.+(2000).+An+Introduction+to+Support+Vector+Machines+and+Other+Kernel-based+Learning+Methods. +London:+Cambridge+University+Press.&ots=xTTd9A_t0f&sig=h5zo1SL2mhMYxWB-99xKF-sJoaQ (accessed on 9 December 2022).
- 63. Morrison, S.; Sosnoff, J.J.; Heffernan, K.S.; Jae, S.Y.; Fernhall, B. Aging, hypertension and physiological tremor: The contribution of the cardioballistic impulse to tremorgenesis in older adults. *J. Neurol. Sci.* **2013**, *326*, 68–74. [CrossRef]
- 64. One-Way ANOVA: Independent Samples: I. Available online: http://vassarstats.net/textbook/ch14pt1.html (accessed on 19 December 2022).
- 65. Tukey, J.W. Comparing Individual Means in the Analysis of Variance. Biometrics 1949, 5, 99–114. [CrossRef]
- 66. Li, C.; Wang, Q. Advanced NIR-II Fluorescence Imaging Technology for In Vivo Precision Tumor Theranostics. *Adv. Ther.* **2019**, 2, 1900053. [CrossRef]
- 67. Meng, X.; Pang, X.; Zhang, K.; Gong, C.; Yang, J.; Dong, H.; Zhang, X. Recent Advances in Near-Infrared-II Fluorescence Imaging for Deep-Tissue Molecular Analysis and Cancer Diagnosis. *Small* **2022**, *18*, 2202035. [CrossRef]
- Wu, H.; Wu, H.; He, Y.; Gan, Z.; Xu, Z.; Zhou, M.; Liu, S.; Liu, H. Synovitis in mice with inflammatory arthritis monitored with quantitative analysis of dynamic contrast-enhanced NIR fluorescence imaging using iRGD-targeted liposomes as fluorescence probes. *Int. J. Nanomed.* 2018, 13, 1841–1850. [CrossRef] [PubMed]
- 69. Zhang, R.; Ma, J. Feature selection for hyperspectral data based on recursive support vector machines. *Int. J. Remote Sens.* 2009, 30, 3669–3677. [CrossRef]
- 70. Khaire, U.M.; Dhanalakshmi, R. Stability of feature selection algorithm: A review. J. King Saud Univ.-Comput. Inf. Sci. 2022, 34, 1060–1073. [CrossRef]
- 71. Brownlee, J. Data Preparation for Machine Learning: Data Cleaning, Feature Selection, and Data Transforms in Python. 2020. Available online: https://books.google.com/books?hl=en&lr=&id=uAPuDwAAQBAJ&oi=fnd&pg=PP1&dq=Data+Preparation+ for+Machine+Learning+Data+Cleaning,+Feature+Selection,+and+Data+Transforms+in+Python&ots=Cl7GwbdUoS&sig=g-7t_iDmR16YA-r3zPDu1pHnQos (accessed on 29 December 2022).
- 72. Alieva, M.; van Rheenen, J.; Broekman, M.L.D. Potential impact of invasive surgical procedures on primary tumor growth and metastasis. *Clin. Exp. Metastasis* **2018**, *35*, 319–331. [CrossRef]
- 73. Lerman, C.; Trock, B.; Rimer, B.K.; Jepson, C.; Brody, D.; Boyce, A. Psychological side effects of breast cancer screening. *Health Psychol.* **1991**, *10*, 259–267. [CrossRef]
- Zhou, R.; Wang, S.; Wen, H.; Wang, M.; Wu, M. The bispecific antibody HB-32, blockade of both VEGF and DLL4 shows potent anti-angiogenic activity in vitro and anti-tumor activity in breast cancer xenograft models. *Exp. Cell Res.* 2019, 380, 141–148. [CrossRef] [PubMed]
- 75. Jimeno, A.; Moore, K.N.; Gordon, M.; Chugh, R.; Diamond, J.R.; Aljumaily, R.; Mendelson, D.; Kapoun, A.M.; Xu, L.; Smith, D.C.; et al. A first-in-human phase 1a study of the bispecific anti-DLL4/anti-VEGF antibody navicixizumab (OMP-305B83) in patients with previously treated solid tumors. *Investig New Drugs* 2019, 37, 461–472. [CrossRef]
- Mohammadlou, M.; Salehi, S.; Baharlou, R. Development of anti DLL4 Nanobody fused to truncated form of Pseudomonas exotoxin: As a novel immunotoxin to inhibit of cell proliferation and neovascularization. *Anal. Biochem.* 2022, 653, 114776. [CrossRef]
- Couch, J.A.; Zhang, G.; Beyer, J.C.; de Zafra, C.L.Z.; Gupta, P.; Kamath, A.V.; Lewin-Koh, N.; Tarrant, J.; Allamneni, K.P.; Cain, G.; et al. Balancing Efficacy and Safety of an Anti-DLL4 Antibody through Pharmacokinetic Modulation. *Clin. Cancer Res.* 2016, 22, 1469–1479. [CrossRef]
- Yan, M.; Callahan, C.A.; Beyer, J.C.; Allamneni, K.P.; Zhang, G.; Ridgway, J.B.; Niessen, K.; Plowman, G.D. Chronic DLL4 blockade induces vascular neoplasms. *Nature* 2010, 463, E6–E7. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.