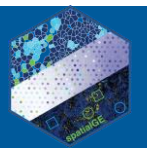
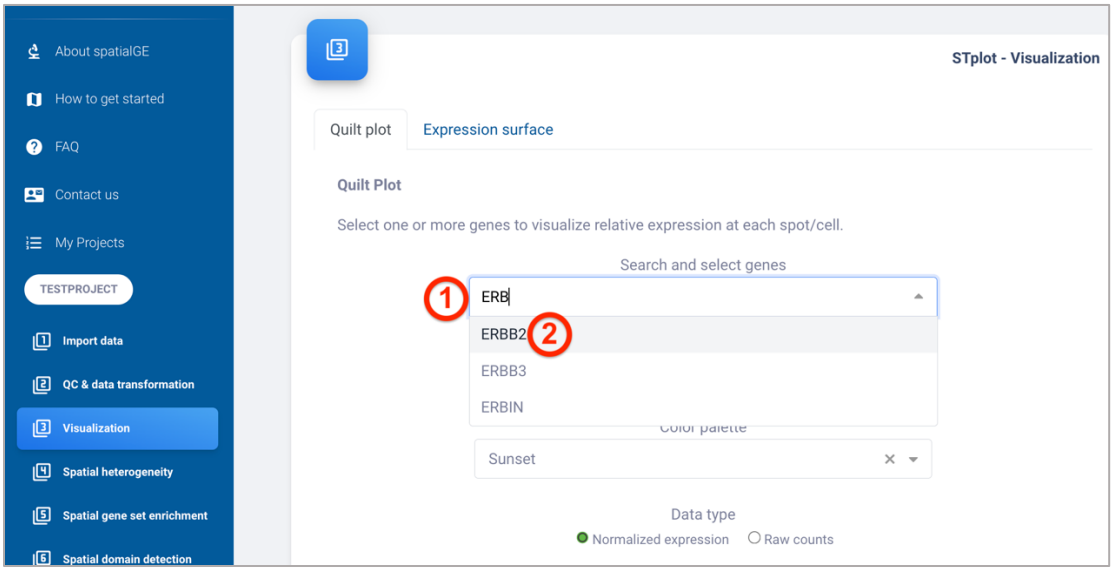


When the normalization procedure (QC & data transformation module) is completed, the **Visualization** module is automatically enabled. The **Visualization** module provides spatialGE users with functionality to generate high-quality figures showing the expression of genes in spatial context. Two options are available: **Quilt plot** and **expression surface**.

Quilt plot

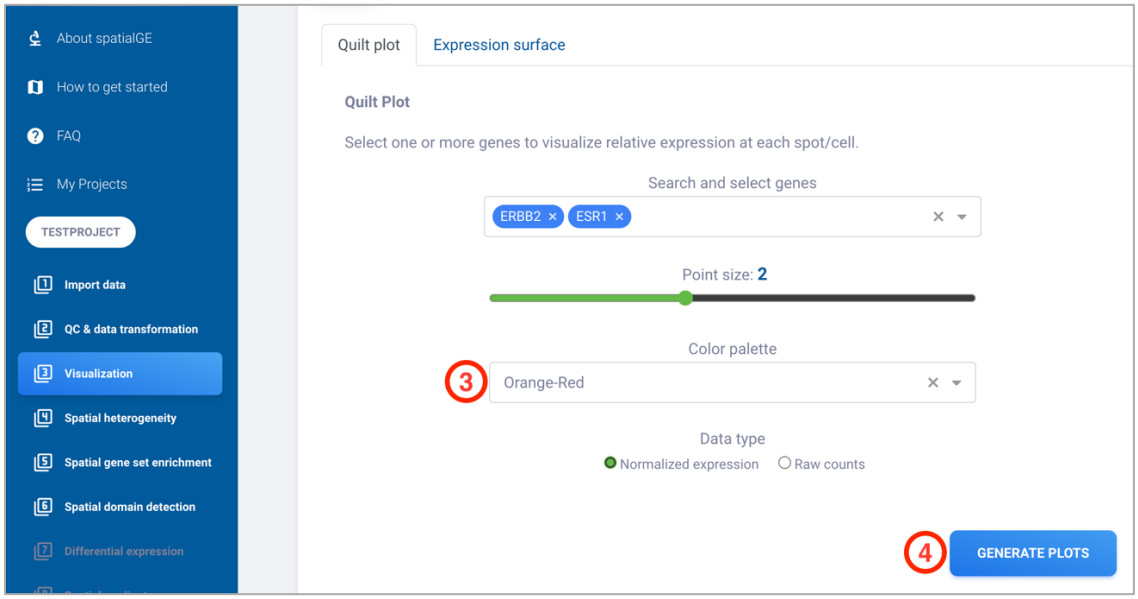
The quilt plots present each Region of Interest (ROI), spot, or cell as a dot, and the dot is colored according to the expression of one or more genes selected by the user. It is a quick way to present spatial gene expression. The interface of **Quilt plot** has been designed to facilitate comparative analysis across samples. To start, please type the names of a few genes in the **Select and search genes** box. For this example, the gene names *ERBB2* (a.k.a. HER2) and *COL1A1* (a collagen gene) will be used. Notice that this textbox allows you to partially input a gene name and displays matching names for your convenience. When the gene name shows in the drop-down list, please click on it.

Note: The **Select and search genes** textbox is case-sensitive, for example ERBB2 is not the same as erbb2.

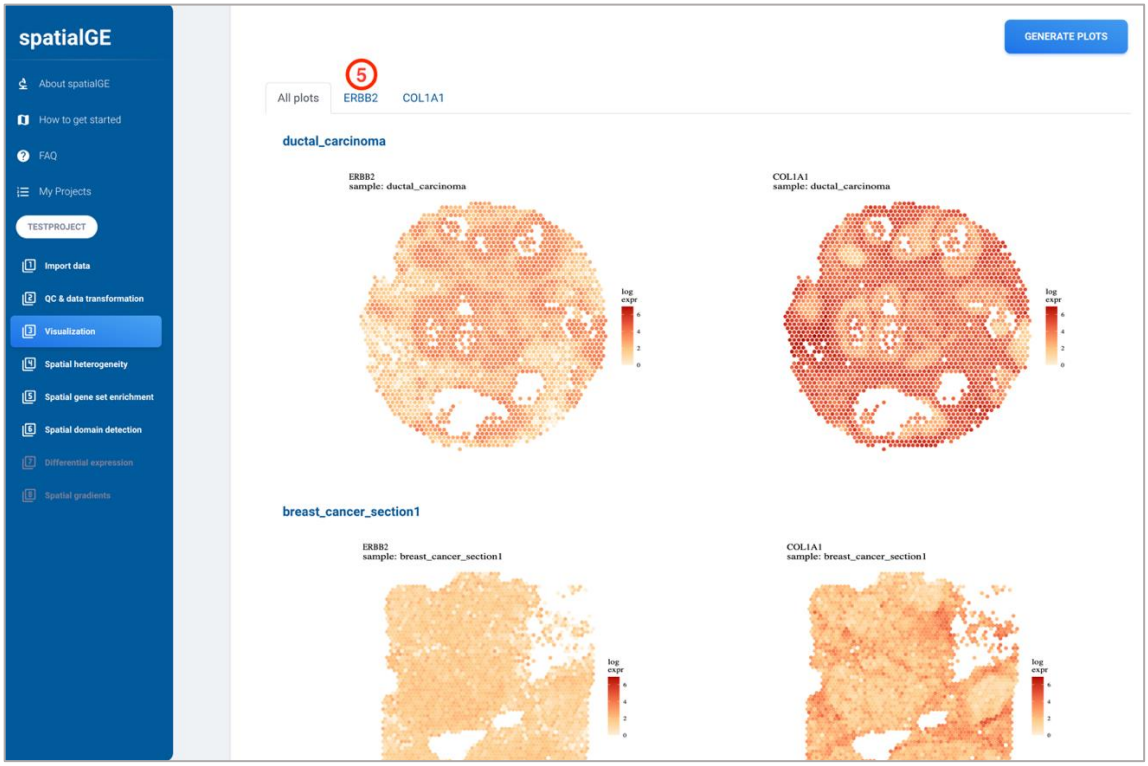


The **Point size** slider will be left in its default value (2), but the user can select smaller or larger dots to represent the ROIs/spots/cells. Users can also choose a **Color palette** name for the expression values at each ROI/spot/cell. For example, please select from the drop-down the "Orange-Red" palette (the palettes are provided by the Khroma and RColorBrewer R packages). The user also has the option to display the normalized or raw counts by selecting the appropriate option in **Data type**. Most users likely will probably want to plot the normalized expression, and for this example, the normalized will be plotted too.

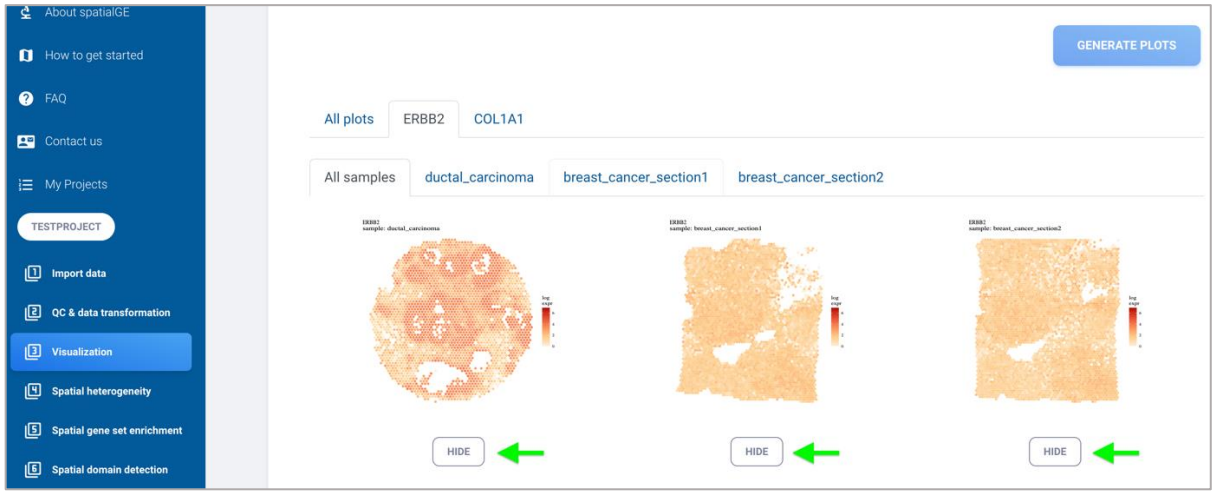
To make the plots, please click **GENERATE PLOTS**.



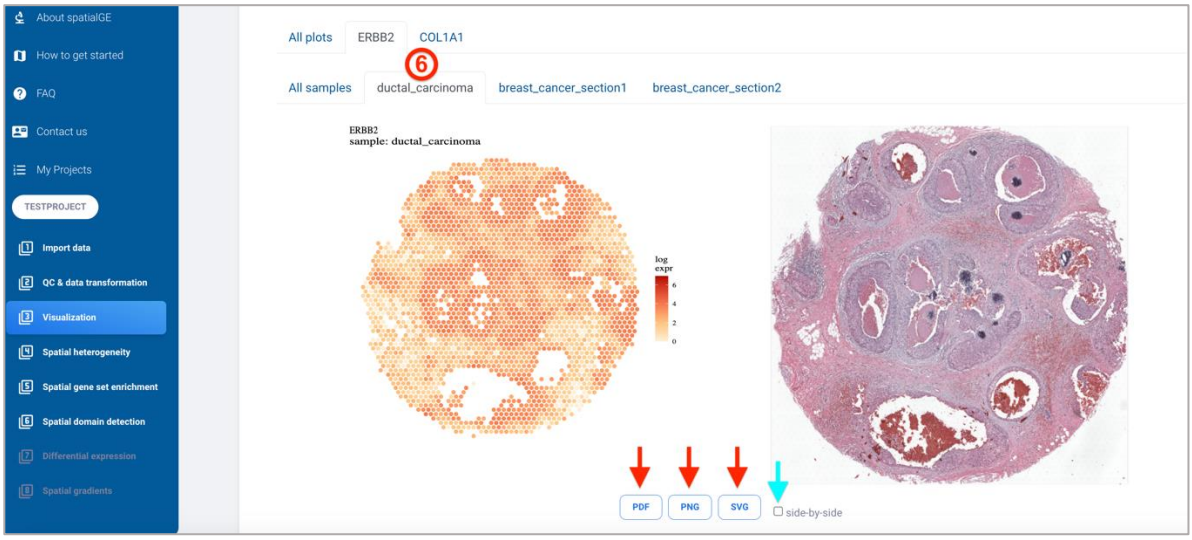
All the plots for all genes and samples are shown in the new **All plots** tab, but you can look at plots for each gene by clicking on one of the adjacent tabs. For example, please click the *ERRB2* tab.



Within each gene tab, you will find tabs for each sample plus an additional tab named **All samples** to facilitate comparative analysis. If the user wishes to hide some of the samples, they can do so by clicking on the **HIDE** buttons (green arrows) under each plot.



Next, please click on the tab with the "ductal_carcinoma" name. This tab allows you to see the quilt plot for a specific gene (*ERBB2* in this case) and specific sample ("ductal_carcinoma" sample). If a tissue image is available, spatialGE displays it here next to the quilt plot. At this point the **PDF**/**PNG**/**SVG** (red arrows) buttons become available to download the image. In addition, by checking the **side-by-side** (turquoise arrow) addition box and then on the button of the desired image file format, the downloaded file will contain the both the quilt plot and tissue image.

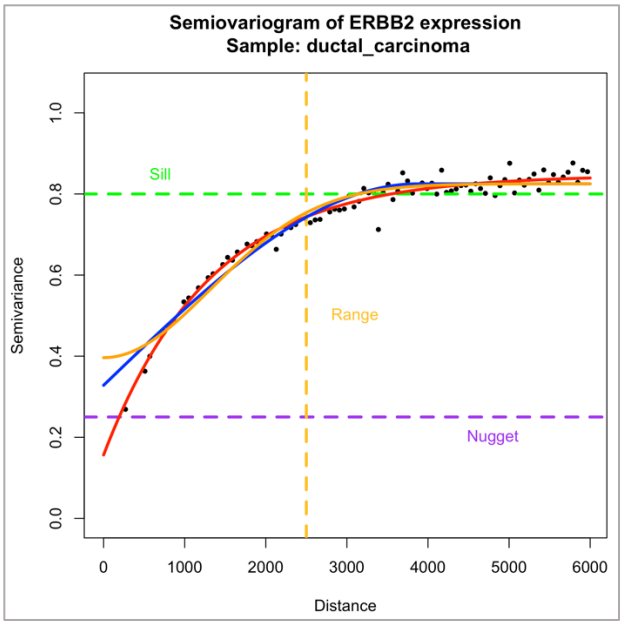


Expression surface plot

Some spatial transcriptomics platforms do not produce data at the single-cell level, and expression is measured at the "mini-bulk" level. Examples of these are the Visium samples in this documentation, in which each spot is approximately 55 microns in diameter, encompassing more than one cell. In addition, "mini-bulk" spatial transcriptomics do not sample each cell within the tissue, but only those captured within the spots. In this case, it is possible to use apply spatial statistics to estimate the expression values of the unsampled areas based

on the sampled areas. This procedure is known as spatial interpolation and creates gene **expression surfaces**. In spatialGE the method known as "kriging" has been adapted to achieve this goal.

The generation of expression surfaces assumes that the expression values of two given ROIs/spots/cells are correlated by their distance. In other words, two spots that are closer are likely to have similar expression values than two spots thar are distant. This correlation between two given spots is often visualized in the form of a semivariogram. Below, a semiovariogram of the expression of *ERBB2* in the "ductal_carcinoma" sample is shown. In this semiovariogram, it can be seen that the expression of *ERBB2* in this spatial transcriptomics sample is correlated to distance, and this correlation can be modeled using an exponential (red line), spherical (blue line), or gaussian (yellow line) covariance function.

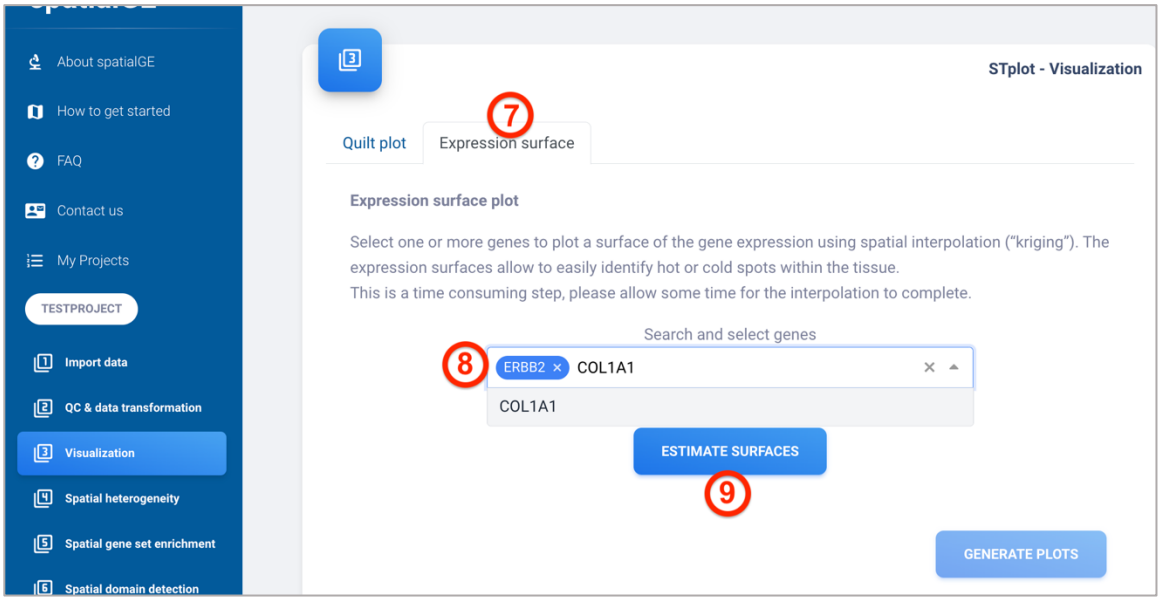


In spatialGE, the type of covariance function used to model the correlation of expression between spots is the exponential function.

To perform spatial interpolation on a set of genes, please click the **Expression surface** tab. Please, begin by specifying a set of genes to generate expression surfaces. The genes to be interpolated are specified in the **Search and select genes** textbox. This textbox features auto-completion, and by typing a letter, a list of matching genes is shown. Please, enter ERBB2 in the **Search and select genes** and click on the gene name. Then enter the COL1A1 gene and click on the name. You are encouraged to add more genes, but please keep in mind that kriging is a computationally intensive method and the more genes you add to the list, the more time it takes to complete.

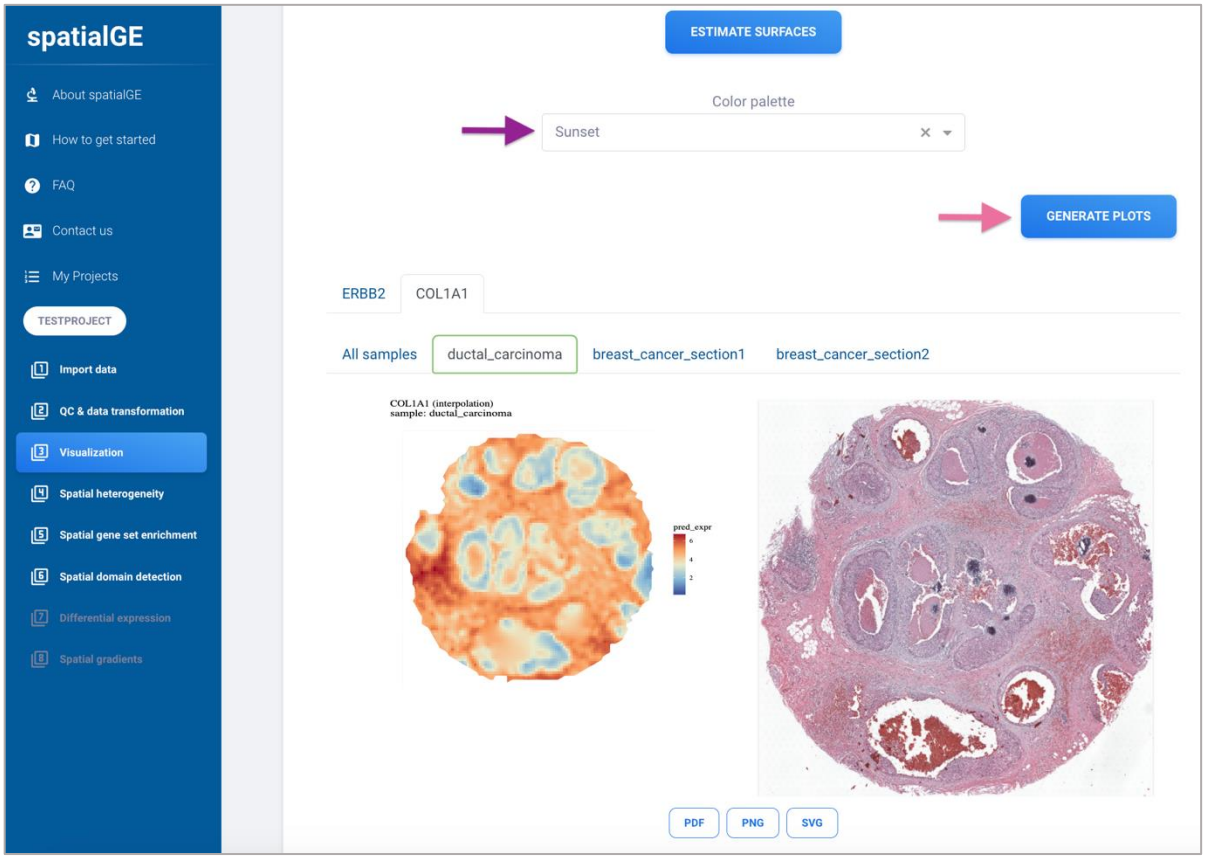
Note: The **Select and search genes** textbox is case-sensitive, for example ERBB2 is not the same as erbb2.

To start the estimation of expression surfaces, please click the **ESTIMATE SURFACES** button.



Immediately the estimation of surfaces is completed, the expression surface plots are showed below. The plots are generated with the "Sunset" color palette, but users can change the colors by selecting one of the options in the **Color palette** drop-down (purple arrow) followed by the **GENERATE PLOTS** button (pink arrow). The palettes are provided by the Khroma and RColorBrewer R packages, please refer to their documentation to learn more about other palettes. Because gene expression is a continuous measure, the diverging and sequential palettes are recommended.

The gene expression surfaces allow for a better exploration of tissue heterogeneity. In the example below, it can be seen that expression of the gene *COL1A1* resembles well the tissue architecture evident in the stained tissue image.



The behavior of the interface is similar as in **Quilt plots**, wherein plots for each gene have a dedicated tab, with sub-tabs for each sample. To display the expression surface next to the tissue image (if available), please go to the sub-tab specific to the sample to be observed.