Table of contents

Dataset (*.lsm, *.czi)

The analysed dataset contains Z-stacked cochlear whole-mount confocal images of gerbil Inner Hair cells (IHC) and Outer Hair cells (OHC) at various characteristic frequencies (CF). The cochleae are stained with four fluorescent antibodies. Phalloidin is used to stain actin, which is frequent in structural elements such as the pillars of the tunnel of corti and the stereocilia of the Hair cells. For fluorescence, 633 nm wavelength laser is used and the emission on the recording is represented with magenta. To stain the ribeye protein which is most frequently found in the IHC nuclei and in ribbon synapses, mouse anti-CtBP2 primary antibody is used as a C-terminus binding protein. A secondary antibody, AF568 Goat Anti-Mouse IgG1 is bound to it with the excitation wavelength of 568nm. On the recording it is represented with red. To find the post-synaptic terminals, a primary antibody, mouse anti-GluR2 is used to stains the glutamate receptors. The secondary antibody, AF488 Goat Anti-Mouse IgG2a is bound to the primary antibody and is excited with the wavelength of 488nm and is represented with green. To stain the myosin in the hair cell bodies, the primary antibody, rabbit anti-Myo7A and the secondary antibody, AF405 Goat anti-rabbit IgG is used. The excitation wavelength is 405 nm and is represented with blue.

The Z-stack images of the four channels are stored using two different extensions, *.lsm and *.czi. Additionally, the files contain the metadata of the recordings which is read and used during the analyses. The following image is acquired using Plan-apochromat 63x/1.3 Oil immersive objective at optical zoom of 1.3x. (Figure 1)

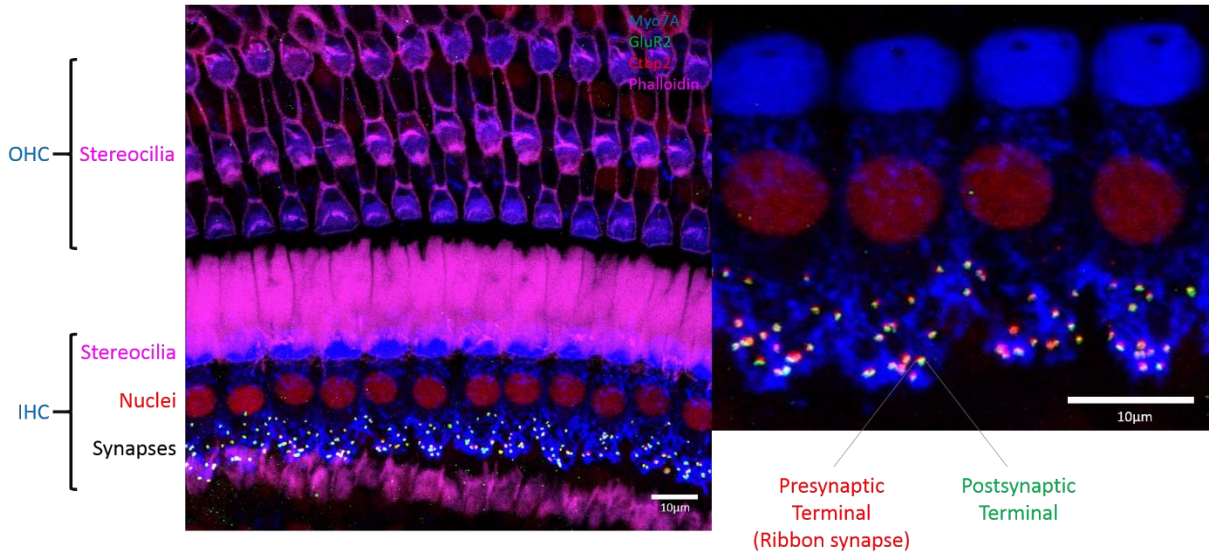

Figure 1 Maximum projection of the four channels of a single recording

Scope of work

Cochlear hair cell loss and synaptic detachment of ribbon synapses are known to play a huge role both in age related and in noise exposed hearing loss. To analyse gerbil cochleae, a confocal microscope was used to record whole-mount images from different CF regions (Figure 1). To quantify the effect of noise exposure and aging on the hair cells and ribbon synapses, an automatic cell and synapse counter is implemented in MATLAB for the recorded whole-mount dataset.

Previously created opensource plugins which quantify synaptic puncta, such as SynapseCounter, which is a plugin for ImageJ (https://github.com/SynPuCo/SynapseCounter), does not support the analysis of stacked images or the extension of saved files. For the analysis of hair cell and synapse count in a single platform, an easy-to-use solution is implemented for the automatic quantification of Z-stacked whole-mount images. A user interface is created in MATLAB AppDesigner, giving an easy way to analyse and manually supervise the results separately for Outer and Inner Hair cells and ribbon synapses.

As the dataset is described above, this description is aiming to give a detailed understanding of the implemented methods used, a brief introduction to the use of the GUI, and finally results of the performance after the manual supervision.

Synapse counting

To count the synapses, two channels are used. The green signal represents the glutamate receptors which are most frequently found on the postsynaptic terminals. The Ribeye proteins in the presynaptic terminals are stained and represented by the red channel. The colocalization of these two markers indicates paired synapses.

As a first step, high intensity noise need to be removed. The intensity of the noise is at least one order of magnitude higher than the highest value of the relevant signal on the recording, and the size is up to two pixels. Dilation and erosion are applied with a vertical line shaped structuring element of two-pixel size to remove the high intensity noise. To use dilation and erosion, a structuring element is created to define the neighbourhood of a pixel. In case of dilation, the value of the pixel is set to the supremum of its neighbourhood pixels defined by the structuring element, while in case of erosion, it is set to the infimum.

For each layer in the Z depth in both channels, two types of threshold were applied. First, a binary image was created for each Z layer with a pixel intensity threshold of 20 %. If the pixel intensity is higher than 20 % of the maximum pixel intensity of the channel, it is set to 1. Subthreshold pixels were set to 0. Based on the binary image, the second threshold utilizes the object size. Objects with the size lower than three or greater than two hundred pixels were set to 0. After the two thresholds, the original pixel intensities were kept for the pixels with nonzero value on the binary image.

On each Z layer on both channels, a moving average is applied with a 3-by-3 square window to smooth the images using stencil processing on GPU. After the moving average, the local maximums are found on each Z layer on both channels. The *(x,y,z)* coordinates and the intensities of the local maximums are stored for further analyses. The stored local maximums are potential synaptic terminals.

After finding the local maximums, false positives are excluded based on vicinity within the channel. If a synaptic terminal is visible on multiple Z layers, the one with the highest intensity value is kept. On different recordings, the actual distance between the Z layers vary, therefore one synaptic terminal can be seen on more or less Z layers. To exclude identical synaptic terminals, a Z distance is defined based on the actual distance (µm) between the Z layers on the recording, as rounding 0.7 divided by the actual distance. Comparing synaptic terminals on neighbouring Z layers in the defined Z distance on the same channel, the synaptic terminals with the highest intensity values were kept if the distance on the *(x,y)* plane between two different synaptic terminal did not exceed five pixels.

Before pairing the synaptic parts, false positives are further excluded based on the vicinity between the channels. The *(x,y)* coordinates of the local maximums of the different channels on the same Z depth within the defined Z distance are compared. Observing a local maximum on one of the channels, if there is no local maximum on the other channel within the defined Z distance within 10 pixel radius on the *(x,y)* plane, the observed local maximum is excluded. This step excludes false positive noise that are not synaptic terminals, and synaptic terminals that are detached from the other synaptic terminal.

The pairing process is performed in three loops. To pair the pre-, and postsynaptic terminals on the different channels, local maximums are compared on the neighbouring Z layers. It is important, that during the pairing the Z distance that was defined earlier is not used. In the first loop, the synaptic terminals are paired that are closer than two pixels on the *(x,y)* plane. Similarly, on the second and third loop, the synaptic terminals closer than five and ten pixels, respectively, are paired. After the

three pairing loops, the *(x,y,z)* coordinates of each synapses are calculated by averaging the paired pre-, and postsynaptic terminals' coordinates.

Inner Hair Cell counting

On the red channel, not only the presynaptic terminals are apparent, but the nuclei of the Inner Hair Cells (IHCs) can also be clearly seen. For counting the IHCs, only the red channel is used.

To remove the signal related to the presynaptic terminals, first, the synapses are found as detailed above. Based on the *(x,y,z)* coordinates of the presynaptic terminal, the pixel intensities are nullified. Nullifying the pixel intensities related to the presynaptic terminals, yields more reliable results in finding the Region of Interest (ROI) of the IHC nuclei. In order to eliminate more signal unrelated to the IHC nuclei, first, erosion then dilation is applied using a disk-shaped structuring element of five pixels as the radius for each Z layers on the red channel. While this process removes the noise unrelated to the IHC nuclei it also weakens the signal of the nuclei.

After nullifying the presynaptic terminals' pixel intensities and removing the unrelated noise on each Z layers, a 2-D maximum projected array is created. On the maximum projected 2-D array, the IHC nuclei are lined horizontally with a slight curvature. Summing up the pixel intensities of each row of the maximum projected array, results in a 1-D array. The location of the maximum value of the 1-D array indicates the location of the middle of the IHC nuclei layer on the *y* axis. (Figure 2)

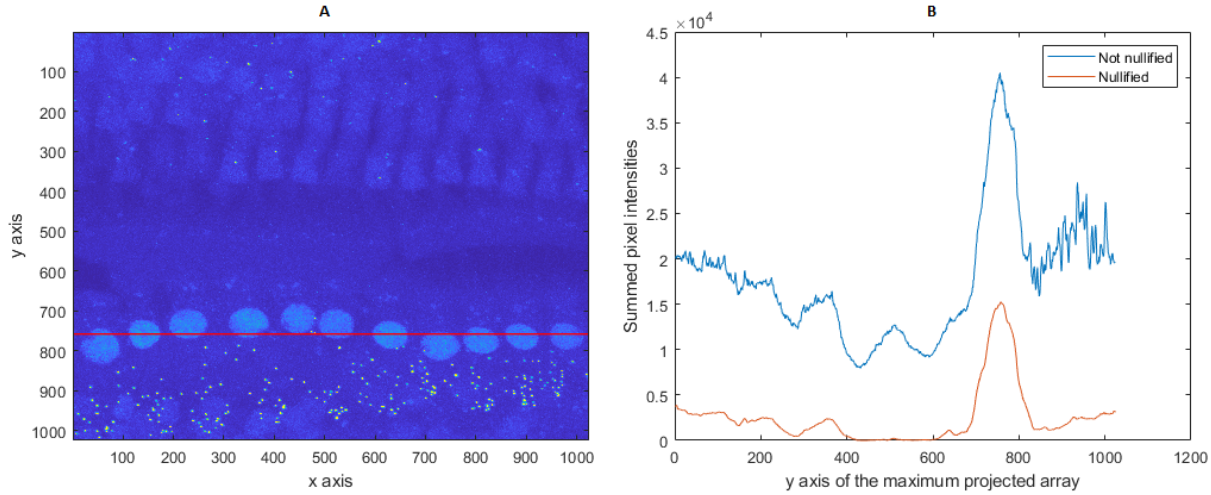

Figure 1 Detection of the middle of the IHC nuclei layer on the y axis. A) Maximum projected image of the original red channel. Horizontal red line indicates the middle of the IHC nuclei layer. B) Red line represents the 1-D array created by summing the rows of the maximum projected 2-D array with nullified presynaptic terminals. Erosion and dilation was also applied. Red line represents the 1-D array created by summing the rows of the maximum projected 2-D array without nullifying the presynaptic terminals. Erosion and dilation was not applied.

To follow the curvature of the IHC nuclei layer and to define the ROI, first, a 40-pixel wide sector is selected around the location of the maximum value on the *y* axis. The 40-pixel wide sector's columns are summed, and the *x* location of the maximum value is chosen. (Figure 3)

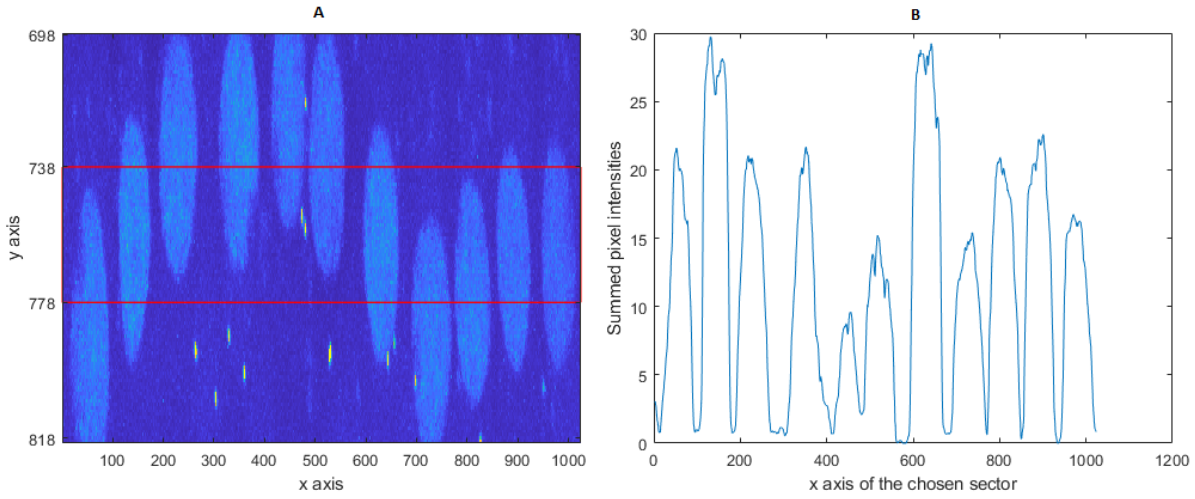

Figure 2 Detection of the location of the maximum x value. A) A part of the maximum projected 2-D array of the original red channel. Red lines are defining the 40-pixel wide sector selected around the indicated y location of the middle of the IHC nuclei layer. B) Summed pixel intensities of the 40-pixel wide sector's columns.

Originating from the chosen *(x,y)* location, the *y* location of the IHC nuclei layer is calculated in both direction along the *x* plane based on the maximum projected 2-D array. To define the curvature of the IHC nuclei region, first, the mean of the pixel intensities are compared of two different squares which are defined based on the chosen *(x,y)* location. The first square is defined as

$$aboveR = [y-20{:}y, x{:}x+50],$$

$$aboveL = [y-20{:}y, x-50{:}x],$$

where *x* and *y* are the chosen coordinates. The second square is defined as

$$belowR = [y{:}y+20, x{:}x+50],$$

$$belowL = [y{:}y+20, x-50{:}x],$$

In case of propagation to the right direction, the mean of the pixel intensities of *aboveR* and *belowR* are compared, and similarly the mean of the pixel intensities of *aboveL* and *belowL* are compared for the propagation to the left direction. Comparing the mean values, the *y* coordinate is adjusted as

$$y' = y + round(\frac{belowM - aboveM}{maxVal} * p),$$

where *aboveM* and *belowM* are the mean of the pixel intensities of the squares defined above, *maxVal* is the maximum pixel intensity value on the red channel and *p* is a parameter which controls the magnitude of the change in the *y* coordinate. The value of *p* is set as 50. In case of propagation to the right direction, the adjusted *y'* value is set as the *y* location of the IHC nuclei layer for the *x* location of *[x+1:x+10]*, and *x* is set as *x+10*. Similarly, in case of propagation to the left direction, the adjusted *y'* value is set as the *y* location of the IHC nuclei layer for the *x* location of *[x-10:x-1]*, and *x* is set as *x-10*. The maximum change in *y* is defined at 10 pixels. Using this propagating approach, the IHC nuclei layer can be defined in a reliable way. (Figure 4)

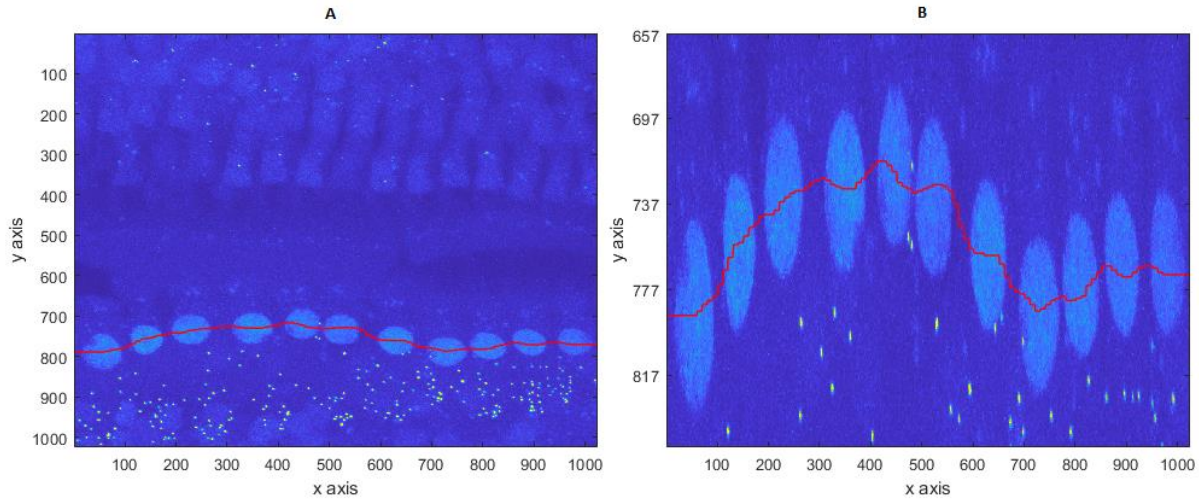

Figure 3 Calculated curvature of the IHC nuclei layer. Red line represents the curvature of the middle of the IHC nuclei layer. A) Maximum projected 2-D array of the red channel. B) A part of the maximum projected 2-D array of the red channel.

The ROI of the IHC nuclei is defined in a sector around the calculated curvature. The width of the sector is defined by the average radius of the IHC nuclei and the real size of a pixel as

$$width = round(\frac{nucleiSize}{pixelSize}),$$

where *nucleiSize* is the diameter of the IHC nuclei defined as 7 µm and *pixelSize* is the real size of a pixel (µm). (Figure 5)

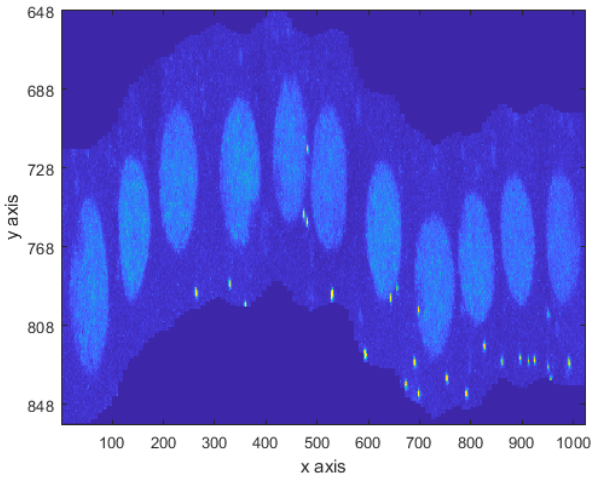

Figure 4 Sector defined around the IHC nuclei layer.

After finding the ROI of the IHC nuclei layer, the original data with the nullified pixels of the pre-synaptic terminals is used. Using erosion first, and then dilation, decreases the possibility to find an IHC nucleus, therefore for further analyses first dilation then erosion is applied. Applying this method after defining the ROI, gives more reliable results on finding the IHC nuclei.

To highlight the relevant information of the ROI, the pixels' intensity of the region is normalized between zero and one, then a modified sigmoid function is applied. The sigmoid function is defined as

$$S(I) = \frac{1}{1+e^{-p*(I-h)}},$$

where *I* is the region with normalised intensities, *p* is the steepness and *h* is the shift of the modified function. The steepness is *p = 14* and the optimal is between 7 and 20. The shift *h = 0,25*, which shifts the function towards the positive direction. The optimal value for the shift is between 0,1 and 0,9.

To smooth the highlighted ROI, a moving average of a round window is applied. The radius is defined as

$$R = nucleiSize/pixelSize/2,$$

where *R* is the radius of the round window, *nucleiSize* is the diameter of the IHC nuclei and *pixelSize* is the real size of a pixel (µm). Before calculating the locations of the IHC nuclei, an arbitrary 60 % threshold is applied on the intensities of the pixels after the moving average.

After the thresholding, the local maximums of the created array of the ROI are found. The *(x,y)* coordinates indicate the centre of the IHC nuclei.

Using the *(x,y)* coordinates of the IHC nuclei and the original red channel, the Z depth of the nuclei can be calculated. To calculate the Z depth of a given IHC nucleus with [X Y] coordinates, a 60-by-60-by-5-pixel *(x,y,z)* cuboid object is used in the original red channel surrounding the [X Y] coordinates of the IHC nucleus, and changing the Z depth. The pixel intensities inside the object are summed for each different Z depths and the results are stored. The *z* coordinate of the IHC nucleus is defined by choosing the Z value related to the maximum of the calculated sums. By calculating the Z depth for each IHC nuclei, the *(x,y,z)* locations of each IHC nuclei are defined.

Outer Hair Cell counting

For counting the Outer Hair Cells (OHCs), the blue and the magenta channels are used. The blue channel represents the myosin in both the OHCs and the IHCs. The magenta channel represent actin, which in case of OHCs, is mostly apparent in the stereocilia. Using the blue and magenta channels the cuticular plate of the OHC can be detected.

First, the ROI must be defined to analyse only the OHC layers and exclude the IHC layer. To find a separating line between the two layers of the different hair cell types, first, the location of the centre of the IHC layer on the *y* axis is calculated based on the red channel as described above. 2-D arrays of maximum projection of the blue and magenta channels are created. Similarly to the case of the red channel, 1-D arrays are created by summing up the pixel intensities of each rows of the maximum projected 2-D arrays of the blue and magenta channel. After normalising the values between 0 and 1 of the created 1-D arrays, a 1-D array, *M*, is created as

$$M = (-1) * BLUE + 1 + MAGENTA\,,$$

where *BLUE* is the normalised 1-D array created from the maximum projection of the blue channel, and *MAGENTA* is the normalised 1-D array created from the maximum projection of the magenta channel. To define the separating line between the OHC and IHC layers, first, the calculated 1-D array, *M*, and the *y* location of the centre of the IHC layer is used. An interval *M'* of *M* of the size of three hundred is defined as

$$M' = M[(Y - 300) : Y]\,,$$

where *Y* is the *y* location of the centre of the IHC layer. The maximum of the array *M'* indicates the *y* location of the separating line between the OHC and IHC layers on the array *M*. (Figure 6)

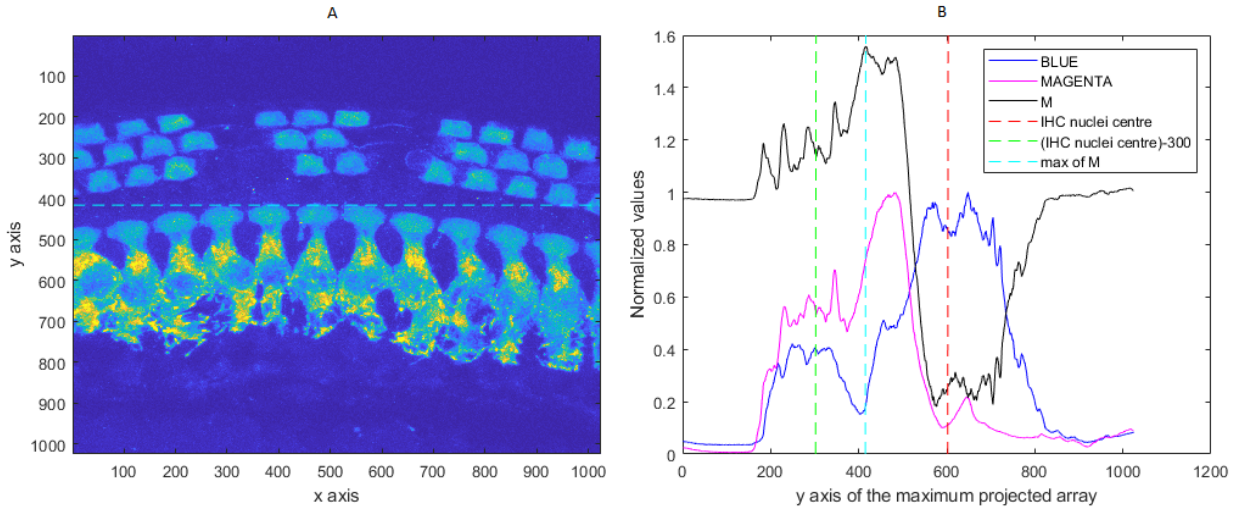

Figure 5 Defining the separating line between the OHC and IHC layers. A) Maximum projected blue channel. Horizontal cyan dashed lines indicate the separating line between the OHC and IHC layers. B) BLUE represents the normalised 1-D array created from the maximum projection of the blue channel. MAGENTA represents the normalised 1-D array created from the maximum projection of the magenta channel. M represents the 1-D array defined as $M = (-1) * BLUE + 1 + MAGENTA$. The vertical blue and red dashed line define the interval M'. The vertical cyan dashed line indicates the location of the separating line.

To take the curvature of the hair cell layers into account, after defining the *y* location of the maximum of the M, similar propagating approach is applied to define the separating line between the OHC and IHC layers as used in case of finding the IHC nuclei layer. The x location is defined as the

location of the maximum of the array created by summing the 40-pixel wide sector's columns around the calculated *y* location. Propagating through the *y* location by comparing the mean values described above in case of the IHC counting, the *y* coordinate is adjusted as

$$y' = y - round(\frac{belowM - aboveM}{maxVal} * p),$$

where *aboveM* and *belowM* are the mean of the pixel intensities of the squares defined in case of IHC counting, *maxVal* is the maximum pixel intensity value on the blue channel and *p* is a parameter which controls the magnitude of the change in the *y* coordinate. (Figure 7)

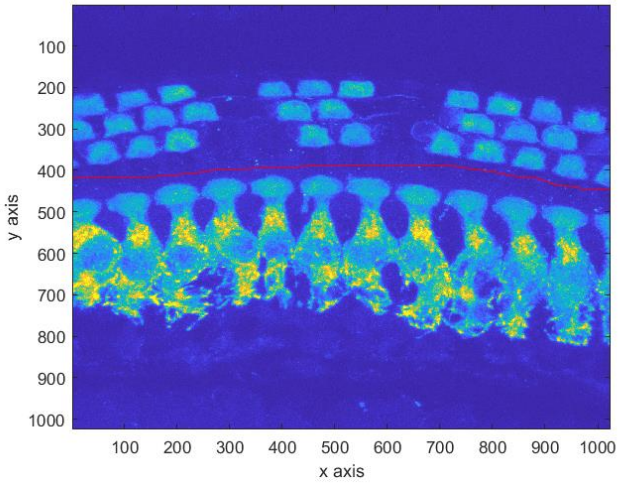

Figure 7 Maximum projected 2-D array of the red channel and the calculated curvature of the separating line between the OHC and the IHC layer. Red line represents the curvature of the separating line.

For noise cancelling on the blue and magenta channels, dilation and erosion are applied with a disk-shaped structuring element of five pixels as the radius on each Z layers.

Separately for both channels, moving averages between the Z layers are applied with a window size of one. The resulting 3-D arrays of the blue and magenta channels are multiplied element-wisely. A maximum projected 2-D array is created from the element-wisely multiplied 3-D array.

On the resulting maximum projected 2-D array, a moving average is applied with a round window of thirteen-pixel radius. The local maximums on the smoothed 2-D array are found, and the *(x,y)* coordinates are stored. The calculated local maximums represent the *(x,y)* locations of the cuticular plates of possible OHCs.

The Z depths of the possible OHCs are defined using a 60-by-60-by-5-pixel cuboid object on the element-wisely multiplied 3-D array, as described above in case of IHC counting. The *(x,y,z)* coordinates of the possible OHCs and the summed intensities in the cuboid object around the possible OHCs are stored.

To exclude false positives, the summed intensities are calculated for each OHCs in the original blue channel in a 30-by-30-by-3 cuboid object around the stored *(x,y,z)* coordinates. Based on the summed intensities in the cuboid objects of the possible OHCs, an arbitrary 40 % threshold is applied to exclude false positives.

Further exclusion of false positive OHCs need to be applied. In many cases, a single OHC is counted multiple times, resulting in false positive OHCs. In most of these cases, the reason of the multiple counting is due to the orientation of the OHC. If the OHC is not oriented perpendicular to the *(x,y)* plane, the local maximum detection on the maximum projected 2-D array will result in false

positive detection. (Figure 8) Many of these false positive OHCs cannot be excluded by intensity thresholding without increasing the number of the false negative OHC. Analysing the area between two possible OHCs (local maximums of *(x,y,z)* coordinates) could reveal if the two compared OHCs belong to different cells or to the same cell.

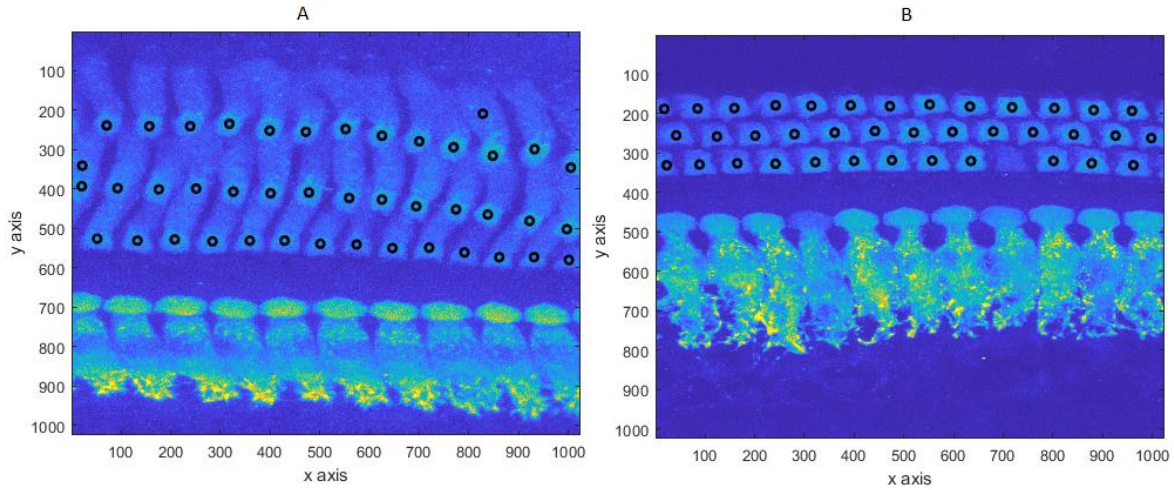

Figure 8 Maximum projections of the blue channels of two different recordings with OHCs oriented perpendicular to the (x,y) plane and oriented not perpendicular. Black circles indicate OHCs. A) OHCs are oriented not perpendicular to the (x,y) plane, causing false positive results. B) OHCs are oriented perpendicular to the (x,y) plane.

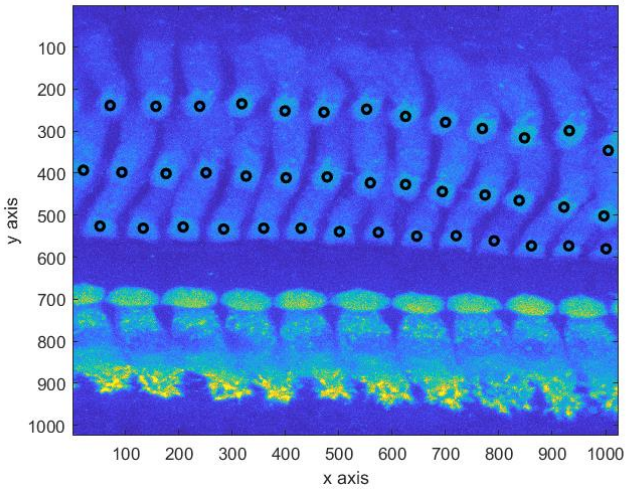

Figure 9 Maximum projection of the blue channel of the same recording as in figure 7 A). Black circles indicate OHCs after the exclusion of false positives due to the orientation of the OHCs.

To exclude duplicate cell counts, local maximums of *(x,y,z)* coordinates closer than 150 pixels on the *(x,y)* plane are compared. As an example, a local maximum at *(100,150,5)* is compared with another local maximum that is found at *(150,100,3)*. A cuboid object of the original blue data is chosen as *(95:155, 95:155, 2:6)*. Maximum projection of the cuboid object is applied on the third dimension, resulting in a 2-D array of the size [60,60]. Analysing the area between the two local maximums on the created 2-D array, can give reliable information whether the two local maximum are originating from two different cells or from a single cell. On the created 2-D array, the pixel intensities are summarised around every pixel in 9-by-9 squares along the line between the two local maximums. The value of the

minimum of these summarised intensities along the line *(I_1)* and the local maximum's summarised intensity with the lower value *(I_2)* are stored and compared. Comparing the I_1 and I_2 can reveal if in the area between the two local maximums cell body is present or not. Using a certain threshold, *Th*, if

$$I_1 < I_2 * Th,$$

the two local maximums originate from two different cells, where *Th* should be between 0.5 and 0.9 empirically. Otherwise, the two local maximums originate from a single cell and the one with the lower intensity value is excluded. (Figure 9)

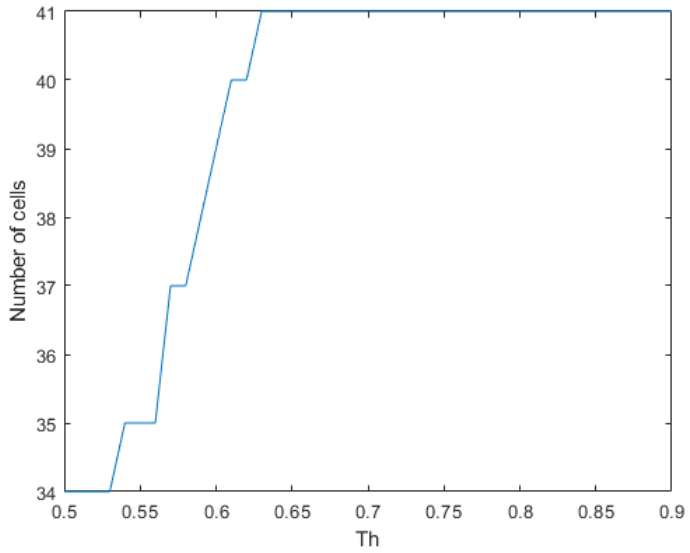

Figure 10 Number of cells at different thresholds (Th) for the recording seen in figure 8 A and figure 9.

The threshold applied, is determined adaptively for each recording. The number of cells after the exclusion is calculated for every *Th* between 0.5 and 0.9 with 0.01 precision. The threshold is used where the cell count is the most stable. (Figure 10)

Graphical User Interface, manual curation

Using MATLAB App Designer, a simple user interface is created for convenient batch processing, and for easy manual supervision and curation.

After opening the main application (SynCounter.mlapp), first, the folder needs to be chosen in which the *.czi and *.lsm files to be processed are found. Using the list box and the radio buttons, one can select to analyse all the files in the folder, the files that have not been analysed or a single file by highlighting it. By ticking the OHC, IHC, Synapse, one can chose which counter to run. The files which have not been analysed are entitled with [NO DATA] in the list box. The results of the analysis are saved in subfolder called *countRes*. (Figure 11.)

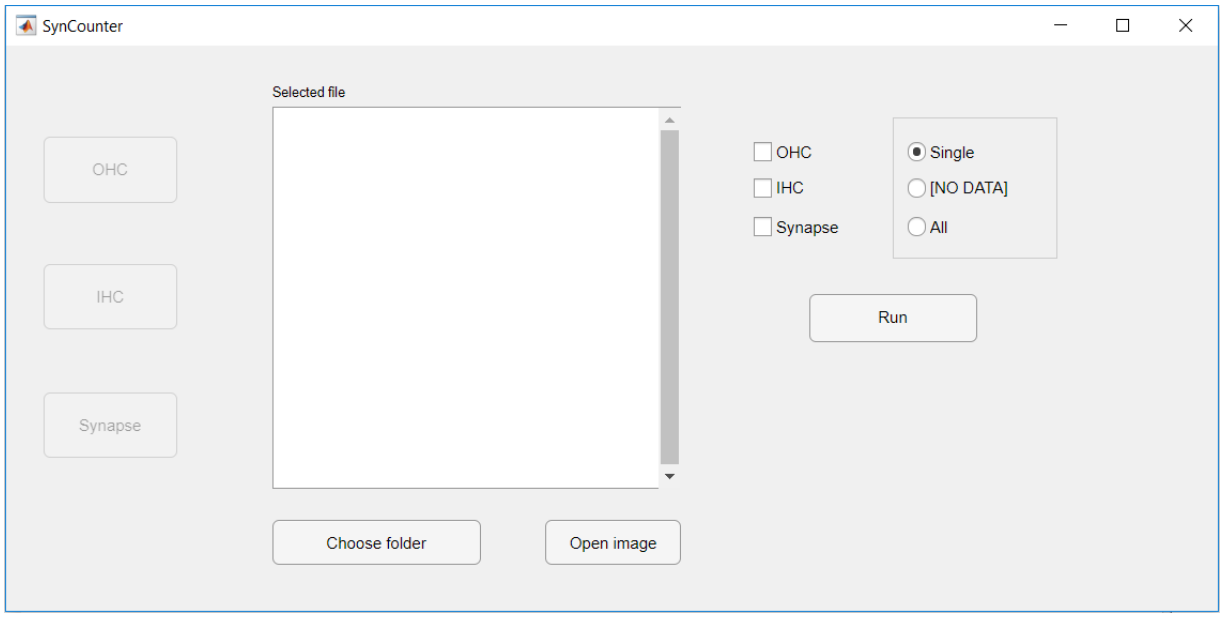

Figure 11 Main application, SynCounter.mlapp

After the analysis, for manual supervision, the analysed and selected image needs to be opened to load the results of the analysis. After opening an image, the appropriate buttons on the left become active to supervise the results by opening a dialogue application. (Figure 11.) Putative cells and synapses are marked with a green circle. In the dialogue application, the figures can be panned and zoomed using the left axes by drag and drop and mouse scroll. Pressing or holding the control button, the right axes will copy the position of the left axes. (Figure 12)

Pressing the left and right arrows, the Z layer can be changed. (Figure 12)

Using the left and right click on the right axes, putative cells and synapses can be added or removed, respectively. Green markers can be excluded by right clicking on the marker on the right axes. This will turn the marker red, which can be included again by right clicking which turns the marker green again. A newly added marker appears in black, which can be removed similarly. (Figure 12)

Supervising the OHC counting, the maximum projection of the blue channel is plotted on the left axes and a single Z-layer can be seen on the right axes. In case of the supervision of the IHC

counting, the red channel is used similarly. Both in case of the IHC and OHC supervision, the putative cells on the right axes are visible only on the calculated Z-layer. (Figure 12)

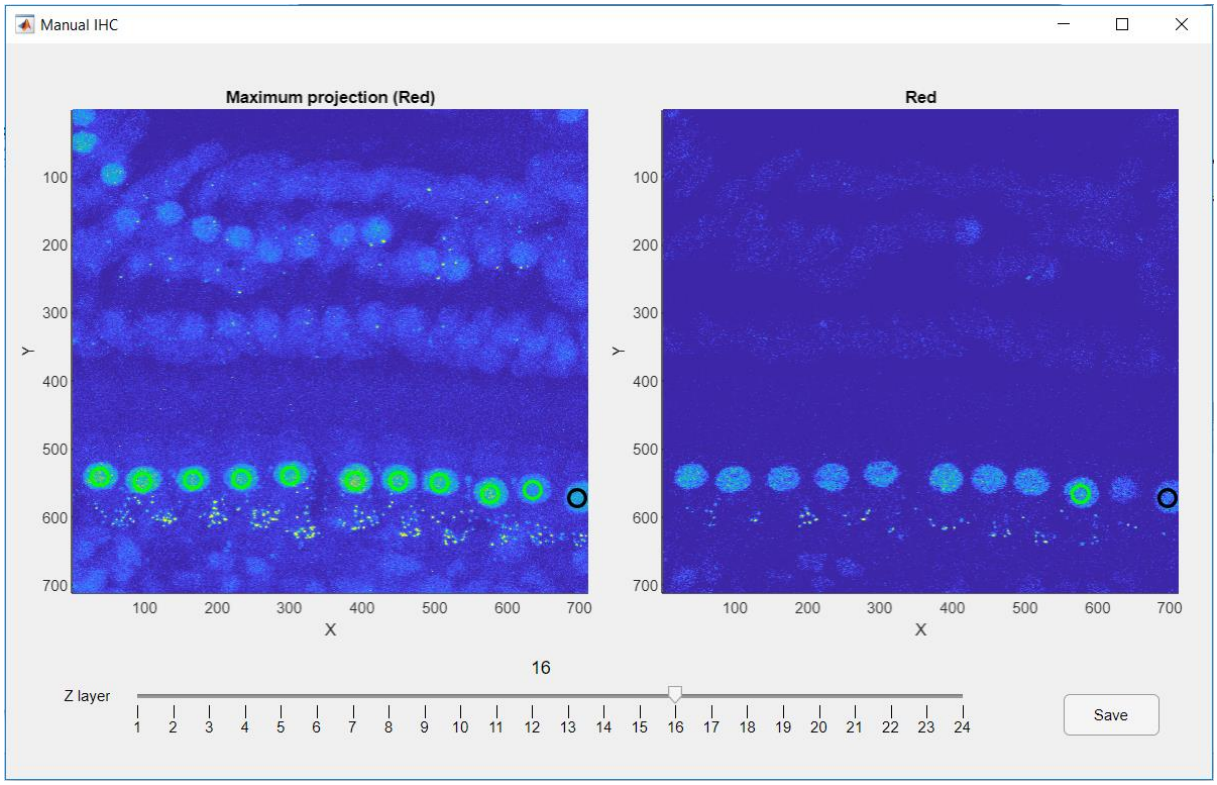

Figure 12 Dialog application for IHC counting. Green circles represent found IHC nuclei, black circle represents manually added IHC nuclei.

Supervising the result of the synapse counting, the red and the green channels are visualised on the different axis. Changing the Z-layer, one marker can be seen either on one or on two layers. If a marker is visible on two neighbouring layers, the synaptic terminals on the different channels have their intensity maximum on different layers. Similarly to the IHC and OHC supervision, figure manipulation can be performed on the left axes, and synapses can be added or removed using the right axes. (Figure 12)

The result of the supervision can be saved, overwriting the originally saved results of the analysis. (Figure 12)

Performance

Due to various reasons, the counters cannot give reliable results in case of certain images. The reasons for the unreliability in most of the cases is tissue deformation. (Figure 12) In case of synapse counting, on some recordings, the green channel is not properly stained, resulting in a noisy image, that the counter cannot overcome. Out of 91 recordings, either due to tissue deformation or weak

staining, the counters have bed performance on 10 (10.99 %) recordings in case of OHC counting, 2 (2.2 %) recordings in case of IHC counting and 8 (8.79 %) recordings in case of synapse counting.

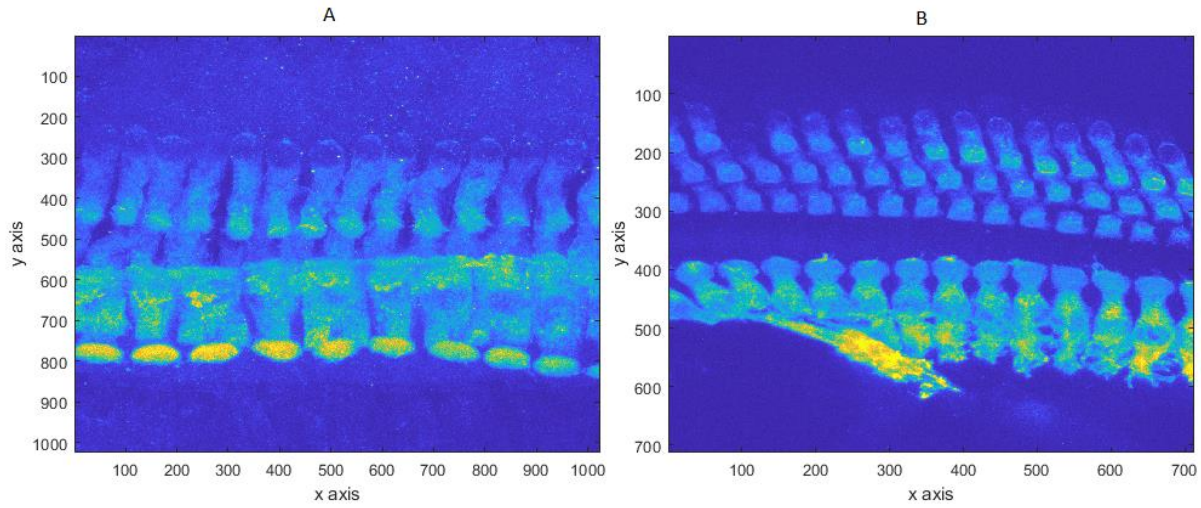

Figure 12 Maximum projection of the blue channel of two different recordings with tissue deformation. A) An example of tissue deformation in the OHC region, resulting unreliable OHC counts. B) An example of tissue deformation in the IHC region, resulting in unreliable synapse count.

On recordings where no tissue deformation is present, the performance of the three counters are tested and the results are supervised manually. A test dataset is chosen quasi randomly, including recording from different CF regions and different file extensions.

In case of OHC counting, 678 OHCs are found. 18 cells were not found and considered as false negatives (FN) and no false positive (FP) appeared. The performance of the OHC counter is 97.41 % accurate.

On the test dataset, 194 IHCs are found out of which 3 are considered as FP and 1 FN is found. The performance is 97.98 %.

As a result of the analysis, 1846 synapses are found out of which 4 are considered as FP and 44 are considered as FN. The performance is 97.46 %.