Comparative Pathology Workbench - an integrated tool for spatial data annotation and analysis with interfaces for visualisation and querying the data

Mike Wicks[1], Michael Glinka[1], Bill Hill[2], Mehran Sharghi[2], Derek Houghton[2], Shahida Din[3], Irene Papatheodorou[4], David Adams[5], Albert Burger[2], Richard Baldock[1], Mark Arends[1]

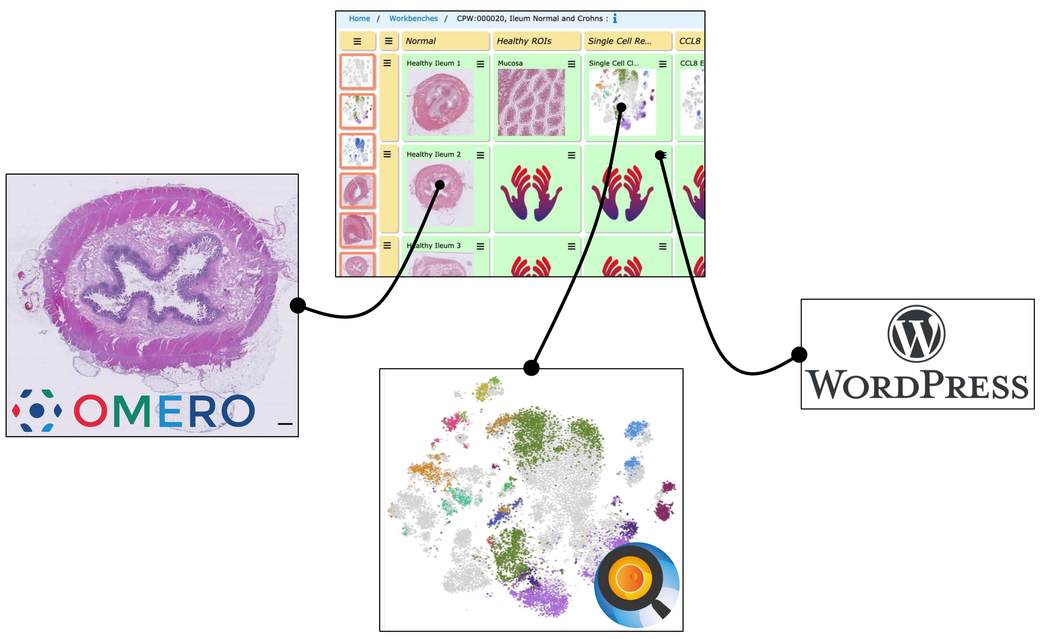
[1]University of Edinburgh, Edinburgh, United Kingdom
[2]Heriot-Watt University, Edinburgh, United Kingdom
[3]NHS Lothian, Edinburgh, United Kingdom
[4]EMBL-EBI, Cambridge, United Kingdom
[5]Wellcome Sanger Institute, Cambridge, United Kingdom

Introduction

The Comparative Pathology Workbench (CPW) is intended as a tool to help pathologists **gather and share related data into a single place**, and to allow them to arrange such data into a meaningful structure for further observation and analysis; no data is stored in the CPW, instead only references to supported data and metadata are stored, within an arrangement as determined by the user. This allows for maximum flexibility for the user, and ensures that as many user styles and preferences are catered for. The CPW uses a **"spreadsheet metaphor" to allow the user to arrange data in a tabular format called a "bench"**. Using such a system allows for easy visualisation of any imaging data to further enhance understanding and sharing of the observations between users.

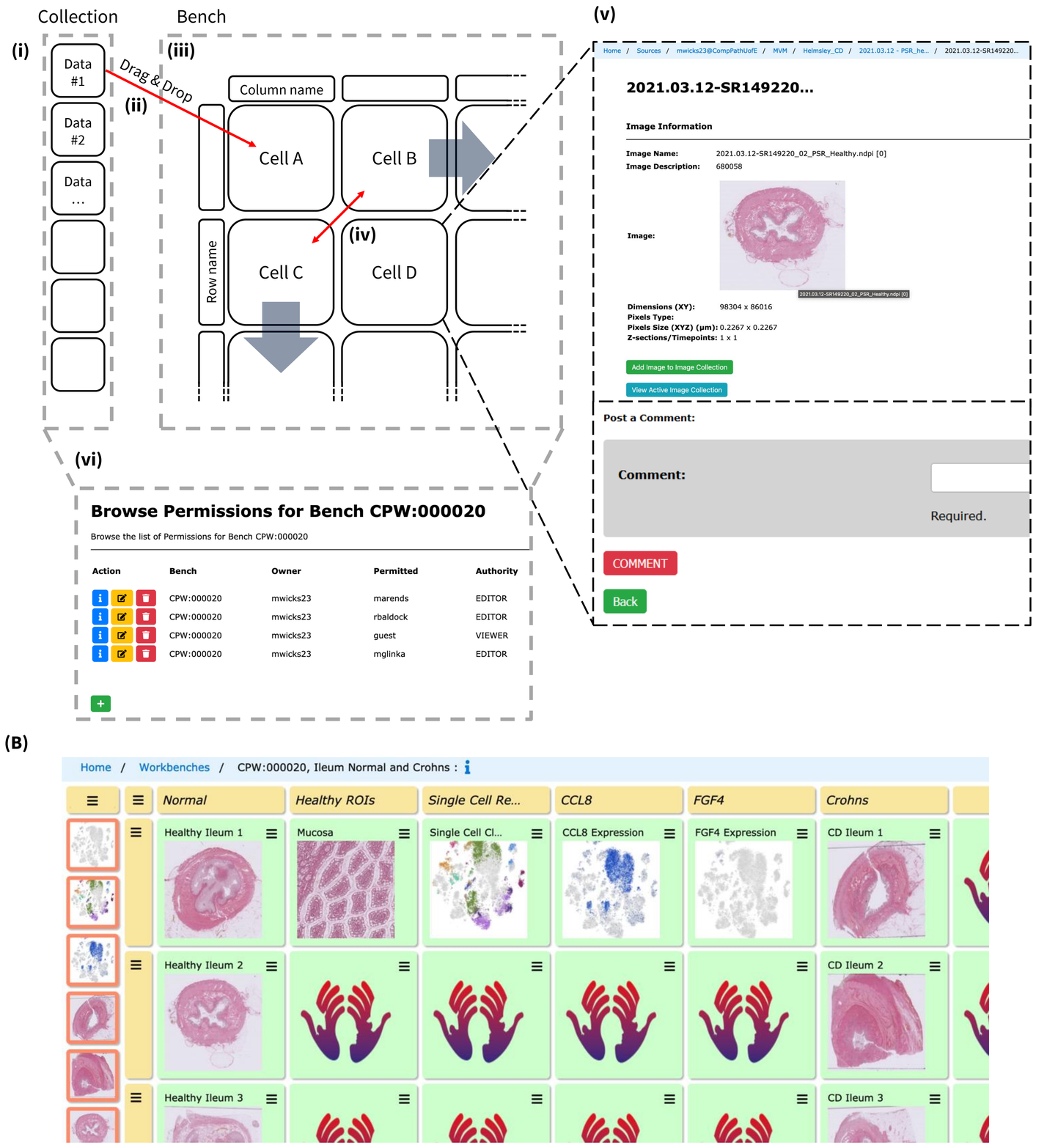
Overview and results

1. Comparative Pathology Workbench (CPW) – overview

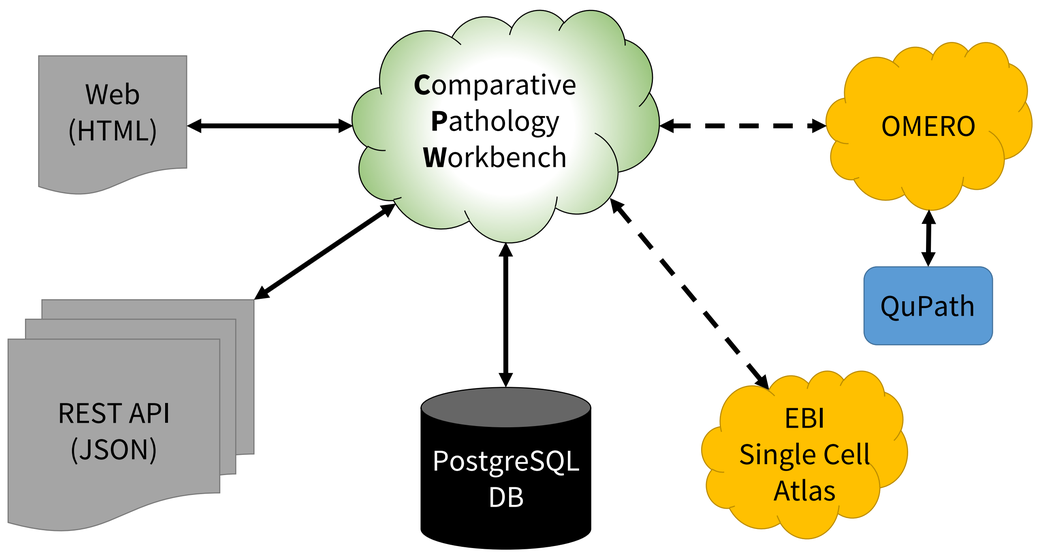


Comparative Pathology Workbench (CPW) – overview of a tool for visualisation of the histological and single cell data. The diagram visualises the flow of the data for visualisation in CPW – a histological and/or single cell data is linked in the tabular view of the workbench. The image data is sourced from the OMERO imaging servers while the single cell data is linked from the EMBL-EBI Single Cell Atlas. Furthermore, the WordPress engine powers the ability to comment and share observations between users.

2. The functionality of the CPW



3. Data exchange



The data exchange diagram of CPW. The diagram shows the data flow and exchange in CPW. No data is stored in the CPW, instead only references to supported data and metadata are stored along with the arrangement of the bench as determined by the user. The CPW uses a small PostgreSQL database to maintain the state of all the benches and collections. It also has a comprehensive REST API, which allows for import and export of benches and collections of images. An integration with the metadata held in EBI's Single Cell Atlas is being developed. This would allow for the inclusion of references to t-SNE plots into bench cells. An integration with analyses generated by QuPath (software for digital pathology image analysis) is also under development. Some QuPath analysis results can now be stored within OMERO, so it is intended to allow the referencing of these results within the CPW.
The clouds represent software systems. Normal arrows represent direct connection and ability to manipulate the data within the systems. Dashed arrows represent links or references to other systems.

(B)

The CPW organises the data in a tabular format, stored in the format of a bench. (A) The diagram shows the functionality of an individual workbench. The references to data/metadata (histological, imaging, single cell, etc.) are linked and stored in a collection (i). The data from the collection can then be dragged and dropped (ii) into the bench in any cell within any rows or columns present in the workbench (iii). Each of the columns or rows can be uniquely named and identified. The cells can be named, re-arranged, swapped or duplicated withing the bench (iv). Individual cells can be viewed to see any additional comments or links to the original data stored in the cell (v). Users can create any number of benches with a set of permissions personalised to each individual user, allowing for viewing, commenting or editing of the bench (vi). (B) Example bench to help visualisation of the diagram.

Conclusions and future work

Comparative Pathology Workbench (CPW) provides a simple to use system to help pathologists and scientists to organise their image and sequencing data plots in order to enhance the analysis of data and to share the observations between users. Its systems were developed to be as user friendly as possible and further development is in progress to add extra functionalities.

Next step involves integration with the metadata held in EMBL-EBI's Single Cell Expression Atlas. This allows for the inclusion of references to t-SNE plots into bench cells, aggregating single cell metadata along with references to histopathological images from OMERO servers. Final coding changes are required to fully support the latest version of the EMBL-EBI SCEA within the CPW.

Further work will involve integration with analyses generated by QuPath, software for digital pathology image analysis. QuPath is now fully integrated with the OMERO software, and is able to generate and share "annotations" to an OMERO server. We will process these annotations and create OMERO "Regions Of Interest" (ROIs), which can be stored as separate images within an OMERO server. These processed images will also be available for referencing within a CPW bench.

Finally, we will integrate access to the Human Cell Atlas Data / Helmsley GCA Data Repository from within the CPW.

Acknowledgements and sponsorship

We would like to thank Helmsley Charitable Trust for funding this project.
We would like to acknowledge work of Michael Cheeseman in aiding development of the project.

GUT CELL ATLAS
An Initiative Supported by the Helmsley Charitable Trust
THE LEONA M. AND HARRY B. HELMSLEY CHARITABLE TRUST
THE UNIVERSITY OF EDINBURGH
HERIOT WATT UNIVERSITY

References: **OMERO:** http://www.openmicroscopy.org/; **EMBL-EBI SCEA:** https://www.ebi.ac.uk/gxa; **QuPath:** https://qupath.github.io/

E:mail: michael.wicks@ed.ac.uk