scientific computing world

Whitepaper

Institute of Cancer Research - Case Study Part I: Beyond the visual: How Vortex speeds up data processing

By Dotmatics

Vortex, the data visualisation and analysis solution from Dotmatics, has reduced the processing time for a Thermal Shift Assay from half a day to less than one hour at the Cancer Research UK Cancer Therapeutics Unit at The Institute of Cancer Research (ICR). When combined with other integrated products in the Dotmatics suite, such as Browser, Vortex can have a great impact on the productivity and efficiency of a research organisation.

Before the software's deployment, many scientists at The Cancer Therapeutics Unit used an Excel spreadsheet to copy and paste raw values from Thermal Shift Assay readers output before evaluating the plots to set the minimum/maximum. They would then copy and paste a subset of that data into GraphPad Prism, before copying the results back into the Excel spreadsheet for further analysis. While this is a common practice it is not very effective as the use of spreadsheets does have its limitations – not least of which is that it is quite laborious.
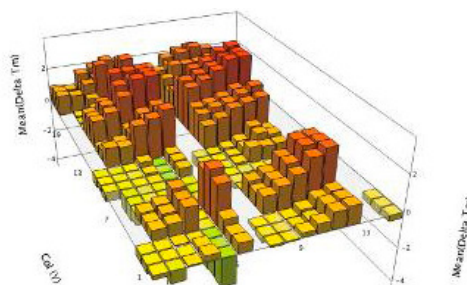
In addition to the inevitable data errors creeping in as a result of copying and pasting, and manually entering data, the half a day taken to process that data had become a bottleneck for the Unit. The solution was to define and implement a standardised method that would eliminate errors and streamline the workflow. By deploying Vortex, each step of the previous process is now handled a standard workflow within one program and completed in less than an hour.

'Most of our users found it easy to adjust to the new methodology of importing the raw data into Vortex and then analysing the displayed results as they were already using it
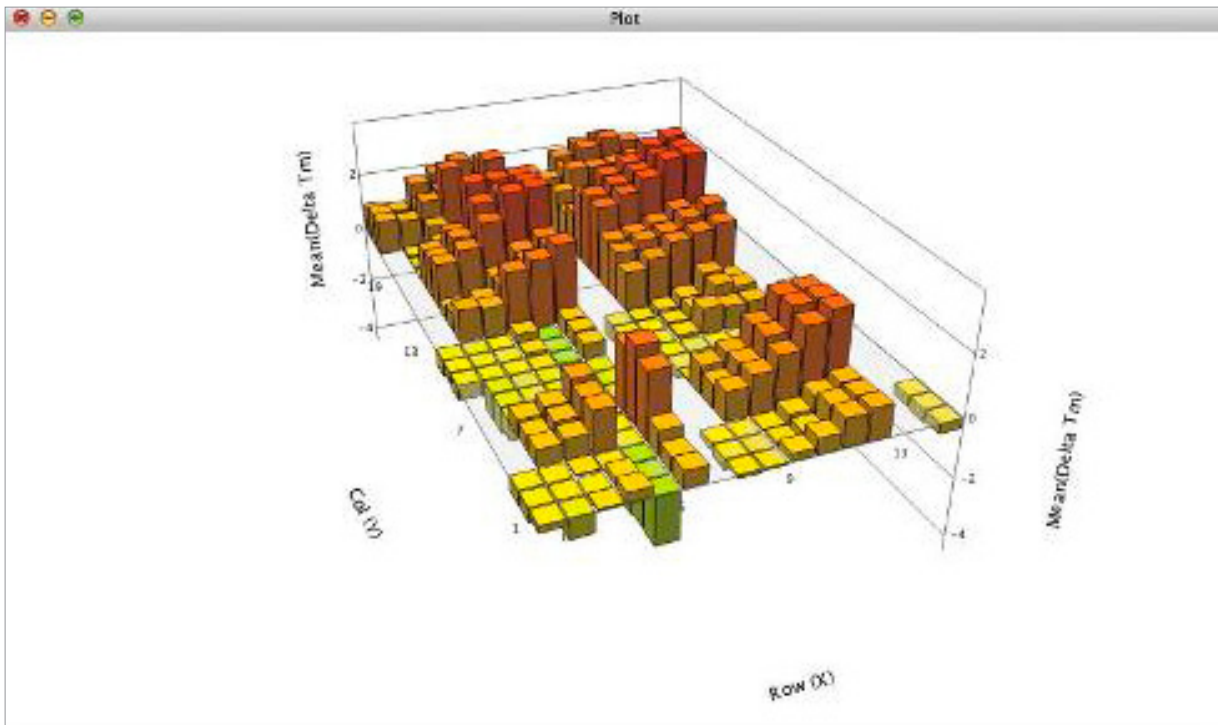
3D Histogram plot showing the thermal shift compared to the controls across the assay plate

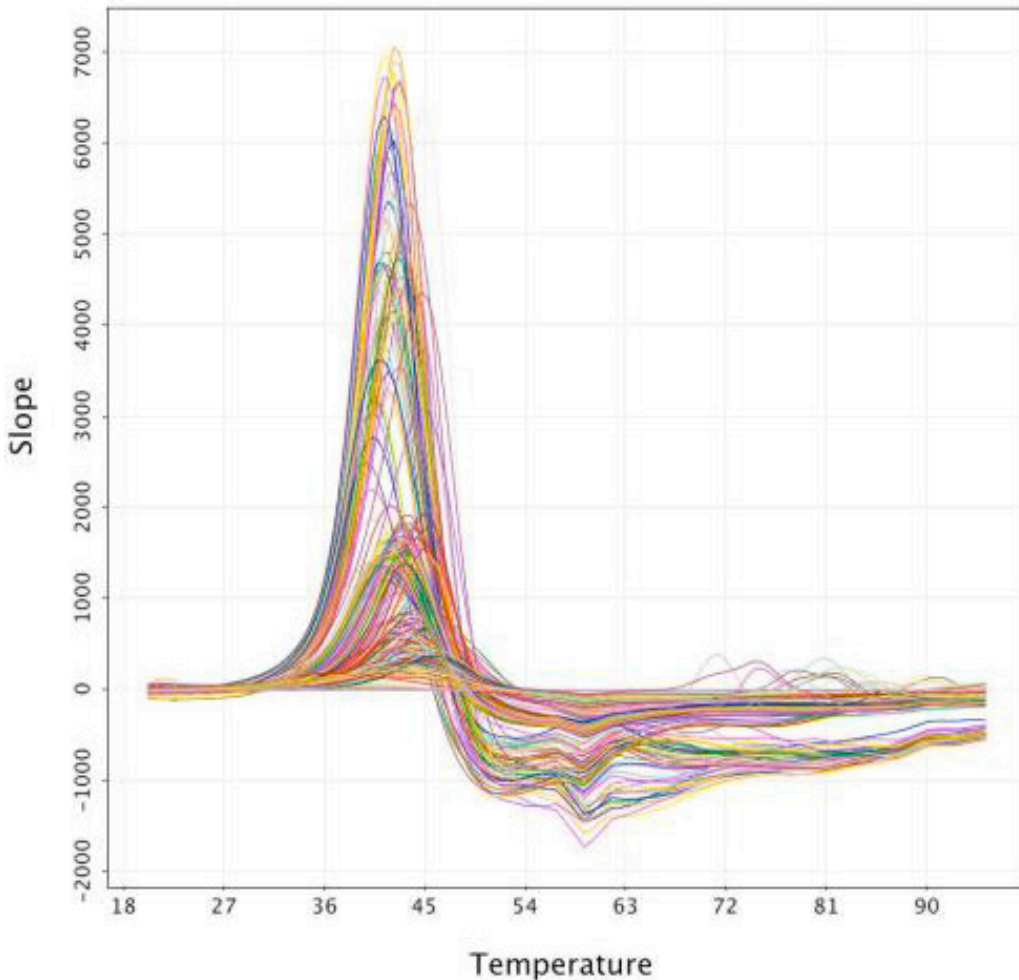*Dr Fiona Jeganathan/Institute of Cancer Research*

for visualising high-throughput screens,' explained Gary Nugent, Higher Scientific Officer, Chemoinformatics at The ICR. 'The scientists who hadn't been exposed to Vortex software before picked it up quite quickly and easily given that they could see how it removed the need to spend a day analysing their plates via Excel spreadsheets. That certainly made it a very attractive proposition for them.'

The Unit soon discovered that a further benefit of moving away from Excel and into something more structured was that the data became far more reproducible. 'Data transformations need to be consistent so that if another scientist were to process the same raw data they would get the same numbers out,' said Nugent. 'Having the workflow written down into a program instead of as a series of instructions written into Excel makes this so much easier and adds to the confidence people have in the work we're doing.'

The Cancer Therapeutics Unit has chosen to centrally provision Vortex, with individual users launching the program via a web start that automatically downloads the latest version and scripts that define best practice data processing. For example, some of these scripts enable users to combine datasets from all their studies and efficiently compare results.

Nugent also noted the benefit of moving from graph fitting the data to using derivative methods to find the point of inflection. The Unit now applies both methods to calculate

the differences. 'Now you can actually see whether or not the shift is real,' he explained. 'Sometimes you can get a shift using the derivative that doesn't necessarily tally up to the shift from a curve fit because there's either a shoulder or multiple peaks in the trace. Being able to differentiate between the two and pick the correct one has been very helpful.'


A plot of the slope of the raw data used as the alternate method of calculating the thermal shifts and identifying multiple peaks and shoulders in the curve

Dr Fiona Jeganathan/Institute of Cancer Research

While Vortex provides comprehensive front end data visualisation and analysis, ICR also integrated R, an open source statistical program, to further expand Vortex functionalities.

After reading a paper where scientists were using R as part of their processing set, Nugent explored the program and found that it was incredibly easy to curve fit and get data back out. 'In order to get the scripting to run I have had to teach myself R, but for anyone who knows how to program in Python, R is not a difficult language to jump to,' he said. The Python scripting functionality in Vortex ties together Vortex and R, and built-in features allowed for quick development of the process and ease of distribution of the resulting analysis
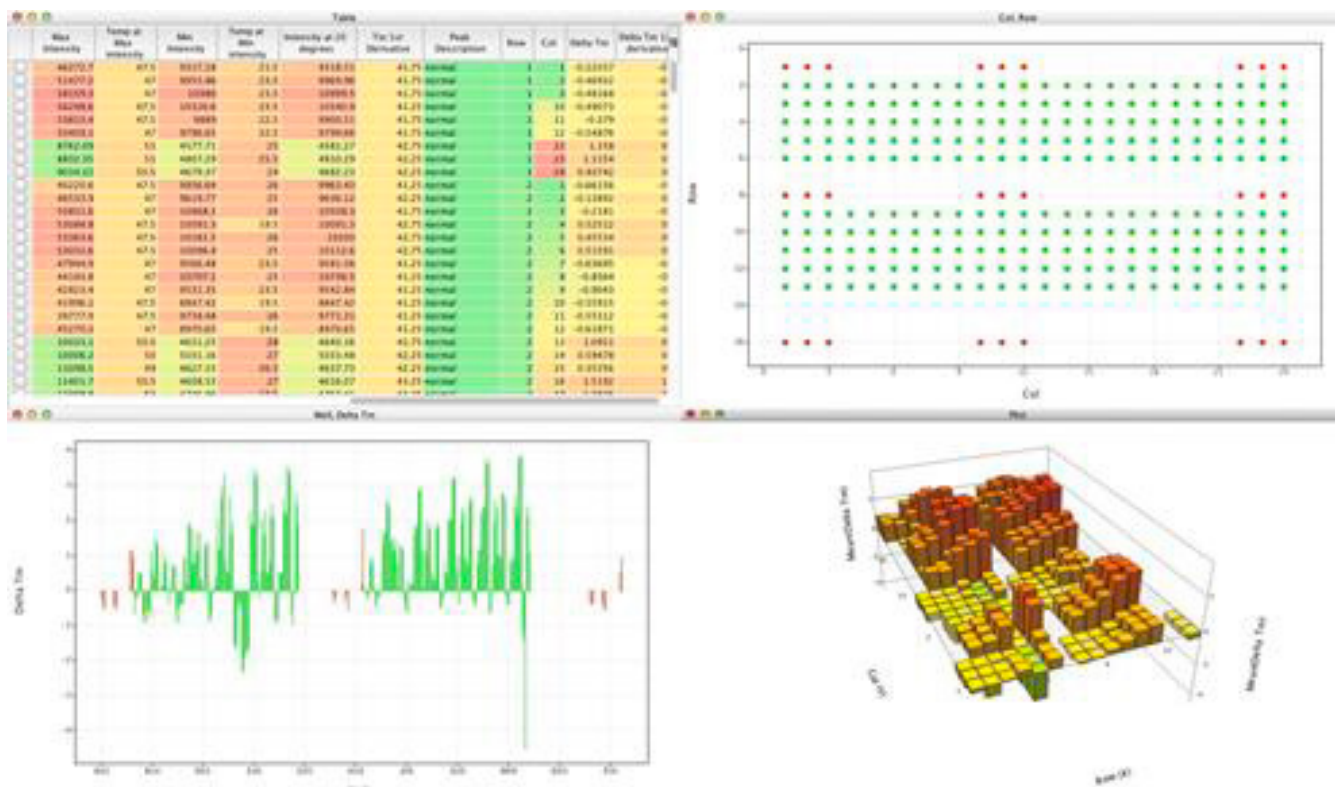
Screenshot showing the output of the processing script. The data table generated by the R script is shown in the top left, the plate map showing the location of the controls and samples in the top right, and at the bottom the plots showing the thermal shifts

Dr Fiona Jeganathan/Institute of Cancer Research

procedures across the team. The connection between the two programs was achieved through the use of examples from the Dotmatics add-on pack and the Dotmatics script share for Vortex. A further resource that helped the process was www.macinchem.org, a website run by Dr Chris Swain at Cambridge MedChem Consulting. The website provides details of many programs, including Vortex, and how best to use them.

'The main thing that has allowed us to develop our internal processes has been the flexibility of Dotmatics as a company,' concluded Nugent. 'The company is very open to exploring different methods and ensuring we can capture and store data in a way that best suits us. Their help really has been invaluable.'