

Three Days of Geeking in Goettingen

Alexander Grosse-Honebrink MSc.



The group of learners from all over Germany and the UK and the tutors from Goettingen.

This summer I was lucky enough to obtain one of the sought after places at the International Workshop for Prokaryotic Genomics and Bioinformatics at the Georg-August University in Goettingen, Germany, and was sponsored by the School of Life Sciences travel grant. It ran from the 15th to the 17th of July and was already the third edition of this excellent course. The programme was, therefore, perfectly routine without it ever becoming boring. I felt that every speaker loved their subject and was very happy to pass on their knowledge and their excitement to us as apprentices to their trade.

The first day had the residents of the Institute for Microbiology and Genomics Dr. Heiko Liesegang and Dr. Andrea Thuermer introducing us to the use of UNIX systems, which are an imperative basis of next generation sequence crunching. After which, we got our hands on some real datasets to learn how to process them from the command line. Later that day, it was on to Dr. Boyke Bunk from the German centre of microbial collections (DSMZ) to show us how easy it is to *de novo* sequence microbial genomes with the 3rd generation sequencing facility PacBio RSII with which DSMZ sequences and finishes up to 24 microbial genomes per week. The introduction was followed by a hands on tutorial in which we casually closed a *Clostridium difficile* genome from PacBio raw data. A piece of cake, really.



Hands on exercises in the computer room of the Institute for Microbiology and Genomics in Goettingen.

The second day gave us an idea of genome annotation which was introduced to us in depth by Professor Alexander Goesmann from the Justus-Liebig-University in Giessen, Germany. A very inspirational talk indeed, which opened my mind to the vast variety of databases and tools used to complete a genome. After this Dr. Anja Poehlein showed us how it is done in real life and introduced us hands-on to the online genome annotation tool IMG (Integrated Microbial Genomes) that can do much more than only annotate and will definitely become a new utensil in my *in silico* tool-belt. After a great lunch in the mensa (restaurant) of the University, Heiko showed us how to use the genome browser Artemis, which we all thought we knew, on a whole new level and how to compare whole genomes and their functions.

The rest of the day was dedicated to an entirely new field for me, RNA-seq. Dr Konrad Foerster of Wuerzburg University gave an energetic talk spiked with some nice statistic basics praising the technology and making it clear at the same time that genes are not the only interesting thing that can be found in a microbial genome. After that Dr Sonja Voget (Goettingen) made a clear point that if we wet-lab workers plan a RNA-seq experiment we should better get in contact with the analysts early on. It was not the first and wouldn't be the last time during the course that we heard: "Garbage in = garbage out" meaning of course: if the source material (DNA, RNA samples) is of bad quality or extracted with the wrong kits the data would be incomplete or missing in the end. We then put the theory to the test and analysed some RNA-seq samples and visualized the results using the in-house developed tool TraV.



The most kissed girl in Goettingen. The Gaenseliesel (goose-girl).

The evening held a few surprises for us with a nice social program. First, we had a tour of the University town of Goettingen with its lovely preserved old town and its 17 real and 45 associated nobel prize winners. We were really tempted to kiss the most kissed girl in town, the Goose-girl (Gaenseliesel) a fountain in the city centre. But we were put off the idea being told it was illegal and the only exception is made for graduating PhDs. The dinner was some of the finest food I had in a while in a restaurant that would only source the best ingredients from the surrounding farm area and, of course, one or two German Pils were not amiss.

The third and last day opened again a new field, the data rich working grounds of metagenomics and metatranscriptomics which Professor Rolf Daniel, the head of the Goettinger Institute of Microbiology and Genetics, introduced before Dr Bernd Wemheuer showed and helped us use the tools of the trade.

The workshop ended with Dr Peter Meinicke diving deep into the algorithmic depths of metagenome analysis and made it clear that big data does not need to be linked to long computation. With some good ideas and a bit of programming magic the huge piles of data can become small and after minutes of calculation you can already have a good insight in the structure of the microbial community of your chosen playground. This was again proven at the work desk and Dr Kathrin Asshauer showed us another neat open source program named "R" to get the right statistics out of our data.

It was a very intense three days of learning and I think everyone of the 14 participants will go home with at least one new tool to simplify our lives. I personally had a great overview over the vast variety of work done in bioinformatics and was infected by the motivation of the speakers and tutors. It also showed me that even though the learning curve is steep, I can enter this field without

having a proper informatics background. I will definitely deepen the knowledge gained in these three days and use most of the tools during the rest of my PhD.

Finally, I want to thank Professor Petra Ehrenreich for organising such a professional and insightful workshop and Professor Rolf Daniel for inviting me to profit from the knowledge of his group.