



Networking / Dissemination Activity Scientific Report Form

Form (Word or PDF file) to be completed and uploaded via the online system within two months of the completion of the event for the following activities: working group meetings, seminars, workshops, symposia, conferences, summer schools, training programmes and specialised courses (graduate-level and continuing education), EUROCORES programme invited sessions at larger conferences, working group meetings (e.g. publication meetings).

a) Summary

On Wednesday most participants and the organisers arrived in Leuven. Several organisers installed in the afternoon all training laptops/network in the meeting room and the course material was transferred onto these laptops. The organisers had a working dinner in the evening to go over the 4-day program a final time and to discuss the final practical arrangements. On Thursday morning all participants had to introduce themselves. This clearly showed we had a very mixed population of participants ranging from computer-dummies to hard-core bio-informaticians, which was the audience we aimed for at the beginning. We tried to make the course as interactive and informal as possible, so questions could be asked whenever the participants had one. Furthermore, all organisers were on standby in the room to help out unexperienced participants. The course started by introducing everyone to the Linux operating system, followed by a general introduction on how to use the command line and the most-used commands to manipulate files/directories, as Linux is the main operating system used for development of NGS tools and algorithms and many of the younger participants have never used a command line before (as DOS was already obsolete by the time they were born). In the afternoon we switched to the R package as this is an easy package to use for data manipulation and data processing of large datasets and not so difficult to learn as a regular programming language, like Perl or Python (which are 4-day courses on its own).

I'm not going in to detail over the rest of the program, but the general idea we had was to provide the participants with all the most-used tools/programs to process raw sequencing data in different ways: "Command line" way (= user has all the control) to "Graphical user interface/Click and run" way (= easier, but user has less control and doesn't always understand what's happening). We deliberately chose for the extremities as we were keeping in mind that participants had different levels of experience. It is up to them to see after the course which programs they prefer to use. We also provided them with study material (e.g. test files) which they could use to practice even more after the course.

On Sunday during lunch we had a feedback session as we wanted to know where we could improve the course for next year, as both the participants and organisers were pleased with how the workshop went. The general conclusion was that the workshop contained a huge amount of knowledge to be processed in 4 days (especially for the computer-dummies), but that it was practically impossible to leave something out (as you are unable to skip a step in the processing pipeline) and making it even longer than 4 days was not preferable in the participants' opinions. If we can find some sponsoring the whole organising team would very much like to organise this workshop on a yearly basis (with a registration fee?) and we would streamline the course material

between the different speakers even more, to make it easier to follow for the participants. Furthermore, in the future we would make one general scheme/flowchart with processing steps and quality steps to be taken (by combining the individual flowcharts some of the speakers had at this course).

b) Final programme of the event

EuroEPINOMICS NGS analysis workshop - Leuven, Belgium (Park Inn by Radisson hotel)

Thursday 13th of March

- 9:00 General introduction Unix/Linux - set-up directories (Roland Krause)
- 9:15 Introduction to the command line (Roland Krause)
Paths, links, autocomplete, shell variants, compression
Root and users
Software installations - package management vs software installation (Samtools, BWA, Vcftools)
Build DeNovoGear
- 10:30 Coffee break
- 10:45 Tutorial session (Roland Krause)
FASTQ FILES, FASTA, (ls, man, grep, cat, cut, make, zcat, gz, tar)
- 13:00 Lunch
- 14:30 Introduction Shell scripting, pipelines - A little bit of R (Roland Krause)
- 15:30 Tutorial
- 16:15 Coffee break
- 16:30 Tutorial continued
- 18:30 Wrap-up
- 19:00 Dinner (Park Inn)

Friday 14th of March

- 9:00 Processing NGS data (Holger Thiele / Kamel Jabbari)
- 10:00 Coffee break
- 10:30 Tutorial (Holger Thiele / Kamel Jabbari / Patrick May)
Different mapping tools
Re-alignment
- 13:00 Lunch
- 14:30 Quality control - Picard tools (Holger Thiele / Kamel Jabbari)
Variant calling (Samtools/GATK)
- 16:15 Coffee break
- 16:30 Tutorial continued
- 19:00 Dinner (De Troubadour)

Saturday 15th of March

- 9:00 NGS data – visualization, coverage. IGV (Silke Appenzeller)
- 9:30 Visualisation in R (Ingo Helbig)
- 10:00 Tutorial session
Coverage plots in R

10:45 Coffee break
11:15 Introduction to annotation/filtering strategies (Ingo Helbig)
11:45 Tutorial Annovar (Ingo Helbig / Silke Appenzeller)
13:00 Lunch
14:30 Introduction to de novo calling + tutorial (Arvid Suls); CNV calling (Kamel Jabbari)
15:30 Introduction on GALAXY (Geert Vandeweyer)
16:15 Coffee break
16:30 Tutorial GALAXY
18:30 Wrap-up
19:00 Dinner (Park Inn)

Sunday 16th of March

9:00 WGS analysis + Ingenuity (Patrick May)
10:00 Tutorial session
10:45 Coffee break
11:15 Tutorial session: VARBANK (Holger Thiele)
12:15 GenomeComb + tutorial (Arvid Suls)
13:00 Lunch
14:30 Feedback and discussion (all)
15:30 Goodbye Leuven

c) Description of the scientific content of the event (abstracts can be provided)

A general strategy (Best Practice/Gold Standard, as far as it exists) to handle, process and analyse NGS data (= huge files) was presented and taught to the participants of the workshop. As we had several experts of different genome centres/research groups we also had open discussions about differences in processing the data and share the pros and contra's with our participants for more insights in the analyses of NGS data. After the course the dissemination continues as all study material used in the course is provided to the participants and all organisers are still happy to answer any questions via mail or Skype. This is important when people start putting the theory into practice on their own data and run into problems. The only way to learn these analyses is to try them on your own data (because every dataset can have its own little problems) and it is crucial that problems at this stage are dealt with.

d) Assessment of the results and impact of the event on the EUROCORES programme.

With this course we disseminated the expertise in NGS data analysis several of us within the EuroEPINOMICS project gained over the years. By doing this many partner will now be able to process NGS data fully or partially, which enlarges the group of researchers able to analyse the huge NGS dataset we have produced within the project. Also, NGS data produced locally will be processed more efficiently in the future, which is beneficial for the project as interesting findings resulting from these analyses are shared with the consortium. Not only will the data be analysed more efficiently, by bringing together bio-informaticians from different genome centres on this workshop we had the opportunity to streamline the different analysis pipeline used locally and

arrange a sharing policy between the different EuroEPINOMICS datasets (e.g. CoGIE and RES), which will definitely help in identifying more genes implicated in epilepsies more efficiently.

e) List of speakers and participants

1. Arvid Suls, VIB-DMG/University of Antwerp, Antwerp, Belgium (IP, RES)
2. Roland Krause, Luxembourg Centre for Systems Biomedicine/University of Luxembourg, Esch-sur-Belval, Luxembourg (IP, CoGIE)
3. Ingo Helbig, University Medical Center Schleswig-Holstein, Kiel Germany (IP, RES)
4. Patrick May, Luxembourg Centre for Systems Biomedicine/University of Luxembourg, Esch-sur-Belval, Luxembourg (IP, CoGIE)
5. Holger Thiele, Cologne Center for Genomics (CCG)/University of Cologne Cologne, Germany (IP, CoGIE)
6. Kamel Jabbari, Cologne Center for Genomics (CCG)/University of Cologne Cologne, Germany (IP, CoGIE)
7. Katja Kobow, University Hospital Erlangen, Erlangen, Germany (IP, EpiGENet)
8. Silke Appenzeller, University Medical Center Schleswig-Holstein, Kiel Germany (IP, RES)
9. Geert Van de Weyer, Center of Medical Genetics, University of Antwerp, Antwerp, Belgium (Invited speaker)
10. Hiltrud Muhle, University Medical Center Schleswig-Holstein, Kiel Germany (IP, RES)
11. Dorota Hoffman, Inst. Of Mother and Child, Warsaw, Poland (IP, RES)
12. Agnieszka Charzewska, Inst. Of Mother and Child, Warsaw, Poland (IP, RES)
13. Eric Fonteneau, Hôpital de la Pitié-Salpêtrière, Paris, France (AP, RES)
14. Johannes Lemke, University of Leipzig, Leipzig, Germany (IP, RES)
15. Sunray Bilgin, Bogazici University, Istanbul, Turkey (IP, RES)
16. Manuela Pendziwiat, University Medical Center Schleswig-Holstein, Kiel Germany (IP, RES)
17. Markus von Deimling, University Medical Center Schleswig-Holstein, Kiel Germany (IP, RES)
18. Davide Mei, University of Florence, Florence, Italy (AP, RES)
19. Lorenzo Tattini, University of Florence, Florence, Italy (AP, RES)
20. Laura Ortega, IIS- Fundación Jiménez Díaz, Madrid, Spain (IP, RES)
21. Rosa Guerrero, IIS- Fundación Jiménez Díaz, Madrid, Spain (IP, RES)
22. Sarah Weckhuysen, VIB-DMG/University of Antwerp, Antwerp, Belgium (IP, RES)
23. Tania Djémié, VIB-DMG/University of Antwerp, Antwerp, Belgium (IP, RES)
24. Katia Hardies, VIB-DMG/University of Antwerp, Antwerp, Belgium (IP, RES)
25. Dennis Lal, Cologne Center for Genomics (CCG)/University of Cologne Cologne, Germany (IP, CoGIE)
26. Eva Reinthaler, Medical University of Vienna, Vienna, Austria (IP, CoGIE)
27. Flip Mulder, UMC Utrecht, Utrecht, The Netherlands (AP, COGIE)
28. Inès Mademan, VIB-DMG/University of Antwerp, Antwerp, Belgium (IP, RES)
29. Jolien Roovers, VIB-DMG/University of Antwerp, Antwerp, Belgium (IP, RES)
30. Janick Mathys, VIB-headquarters, Gent, Belgium (for ICT support, training laptops)