

Course reflections

K. ALEXA LANG

January 6, 2025

Contents

1	My expectations	2
2	Reflection on attended conferences	2
2.1	Spatial Transcriptomics - Irepan Salvador Martinez	2
2.2	Learning from cancer genomes - Claudia Arnedo Pac	3
3	Final discussion: What did I learn?	4

1 My expectations

As an Erasmus student I was free to choose which courses I wanted to take here at UAB. Therefore I spent quite a lot of time thinking about what I expected from each course in order to pick the best ones for me. It was clear to me that I wanted to do courses on bioinformatics, since a course I attended last semester sparked my interest in the topic. This course was about r analysis of high throughput sequencing. To see how those raw datasets were transformed into usable conclusions and graphics fascinated me.

From the TAB course I hoped to gain a deeper insight on two distinctive topics. Firstly I wanted to acquire a basic knowledge of bioinformatics and additionally I wanted to explore if I could see myself focusing on bioinformatics in the future.

In terms of practical knowledge I hoped to gain basics in python and bash, which I had not previously used. I also wanted to expand my repertoire and revise what I already knew in r. Especially python was very interesting to me since I know that it is often applied in biological research and did not learn it at my home university. As your course contained practicals I hoped for hands on training and a chance to apply and thereby truly understand the discussed topics.

From the seminar I mainly hoped to gain inspiration and an overview of the current opportunities and most important topics in the field. Ideally I wanted to find topics and institutions which would interest me for internships or a master project.

2 Reflection on attended conferences

2.1 Spatial Transcriptomics - Irepan Salvador Martinez

Different from most other seminars, the discussion after and during Irepan Salvador Martinez seminar on spatial transcriptomics focused mostly on the experimental method. After having heard a lot of career advice already it was really fascinating to dive deep into a previously unknown concept. Of course I had heard the term "spatial transcriptomics" before and had the superficial knowledge that it is an advanced form of single cell RNA analysis which contains information on where each cell was localized in the tissue sample. However beyond that I knew nothing about this method and had no idea by which method spatial sequencing was possible.

Spatial transcriptomics is the newest advancement in the field of RNA sequencing and high hopes are put into its impact on research. This was honored for example by the nature magazine electing it as the method of the year 2020. It thereby follows in the footsteps of its predecessor single cell sequencing, which had won the title in 2013. In RNA sequencing a clear trend towards higher resolution and differentiation of distinctive cells is visible in the evolution of methods from bulk to single cell and now spatial transcriptomics.

Two main commercial methods are currently used, both by the company 10x: Visum and Xenium. The main methodological difference lies in Xenium being an in-situ analysis whilst Visum breaks up the tissue sample after bar coding RNA according to its localization in the tissue.

The seminar focused on Xenium, which is based on padlock probe rolling circle amplification. A padlock is an oligonucleotide designed to bind to a specific sequence in a circular fashion. Each padlock contains binding regions for their sequence on each end as well as a specific barcode for recognition. Only if the padlock binds its sequence, ligation into circular form takes place. All circular padlocks are amplified by rolling circle amplification before the tissue sample is analyzed.

Analysis takes place via successive rounds of fluorescent probe hybridization, imaging, and washing.

Different fluorescence markers are used for each barcode, leading to each distinctive fluorescent signal standing for one gene. This principle of analysis of fluorescent signals is known from other methods such as sequencing by synthesis.

Since this method is still very much under development, some obvious limitations remain. Firstly, only previously selected genes for which a padlock is designed and added are analyzed. This need for pre-selection of typically 100-200 genes leads to strongly biased results. Keeping in mind that even 200 padlocks would represent $\leq 1\%$ of the expected 20.000 – 25.000 genes in the human genome, it marks the method as currently only suitable for very targeted research questions as opposed to gaining a general overview of genes transcribed in a tissue. A further limitation is that RNA sequences are not decoded. In case of mutations or changes of the RNA, the padlock would either not bind at all, leading to a false negative or bind normally and information of the mutation would be lost.

10x is currently working on overcoming these limitations by increasing the number of padlocks that can be used in a single analysis. New and improved machines launched last year are already supposed to handle the analysis of up to 5.000 different genes.

This topic had fascinated me especially in showing how interlinked new laboratory methods and bioinformatics are. Developments in both fields have strong synergistic effects and there is a need of both to improve together. For this endeavor people educated in both molecular biology and informatics are needed, which makes me hopeful to find my niche in this space some day.

2.2 Learning from cancer genomes - Claudia Arnedo Pac

Claudia Arnedo Pacs seminar talk stayed with me not only due to my general interest in cancer research but also due to her rich career advice and motivational spirit.

Her research focuses mostly on the identification of driver mutations in cancer cells. Among the many somatic mutations occurring in carcinogenesis, these are the ones driving the disease forward by providing key selective advantages to mutated cells. Their identification enables research in precision medicine, in which prevention, diagnosis and treatment measures are adapted to the unique mutational profile of each patient. Precision medicine has the potential to not only improve the efficacy of treatment but also minimize off target effects which cause devastating side effects in patients.

As stated before, a part of the seminar was dedicated to bioinformatics as a possible future career. Since Claudia Arnedo Pacs has started her career in a wet lab like myself, it was very interesting to learn about her career path. She described the feeling of wanting to truly understand and perform the needed analysis for her bachelors project and therefore getting into bioinformatics. I empathized a lot with this since I have been confronted with the same feeling during my masters studies; mostly generating experimental data in the lab whilst lacking the skills and resources to correctly analyze them myself deeply frustrated me.

Claudia Arnedo Pacs described bioinformatics as consisting of three key areas, ordered in descending importance: Biomedicine, Programming/Computer Science and Statics/Math. In addition to knowledge in these areas, one should bring and/or build a set of different skills. Technical and research skills such as programming, data analysis, visualization and creative critical thinking are needed next to communication (presentation, writing & collaboration) and soft skills (project management, conflict resolution, teaching).

Even if those skills are the ones currently needed, keeping a "learning mindset" is crucial. Depending on career level and work environment, different soft and communication skills are needed. The required research and programming skills are also bound to change consistently with the development of new bioinformatic tools and research methods. However this can not be seen as a constant burden, quite the opposite is true since this leads to keeping critical thinking and self education alive. It is the constant

improvement and changes in the field which turn it into a highly interesting working area.

Claudia Arnedo Pacs also reassured us that the bioinformatics community provides a large amount of resources to learn and train new techniques, understand new fields and connect with others. She encouraged us to use available online resources to learn needed skills such as python programming or statistics basics. Her tips and perspectives motivated me a lot and gave me an idea on how to prepare myself for potentially working in this field.

Additionally to needed skills, potential models of employment were discussed. The three key employers are academic institutions, hospitals and industry. As a student at university, jobs in academic institutions were relatively familiar. However I had spent little thought on working in a hospital and was not really aware of this opportunity. Being shown opportunities which I had not known of reassured me, since it suggests that there is a vast amount of potential career paths, so many that a lot of them we might have never heard of.

3 Final discussion: What did I learn?

Overall the course helped me a lot, both on the technical skills level and the open questions regarding my future career.

The format of doing exercises on our own in organized scripts like R markdowns or jupyter notebooks was very helpful since it allowed us to code ourselves whilst having background information and detailed how-tos available in an intuitive format. Inspired by this, I also started to use these document types for other courses and personal work.

Not only the material but also the structure of the course was very well organized. Objectives and expectations were communicated clearly and in an executable manner for students.

The course felt like a great start to dive deeper into the topics which I wanted to learn anyways, as suggested by Jose I therefore started to practice python using the rosalind project and other online tutorials. Some of the generated material are uploaded to our github in the "extra" repository.

Regarding the format of the course I was surprised but not bothered by how much of our grade comes from group work. Since I have seen the same at many other courses at UAB I believe this is a general difference in university culture between Germany and Spain.

The seminars were instrumental for me in gaining perspective on current topics and opportunities in bioinformatics. I used this in my search for internships and was able to get two internships in the field of bioinformatics for this year. One at a research group in neurological spatial transcriptomics at the Max Planck Institute in Cologne and one in the university hospital cologne in a project identifying drivers of chemotherapy resistance in cancer.

On a more personal note, this course was also one which made me feel very welcome and cared for as an international student. I highly appreciate the way both of you, Martha and Jose, checked in and asked how we were doing and even asking if our other courses were going okay. Besides this I felt very integrated since the everyone always translated to english as much as possible to make it easy on us. All in all I am very happy to have chosen this course and hope to hear again from you in the future!