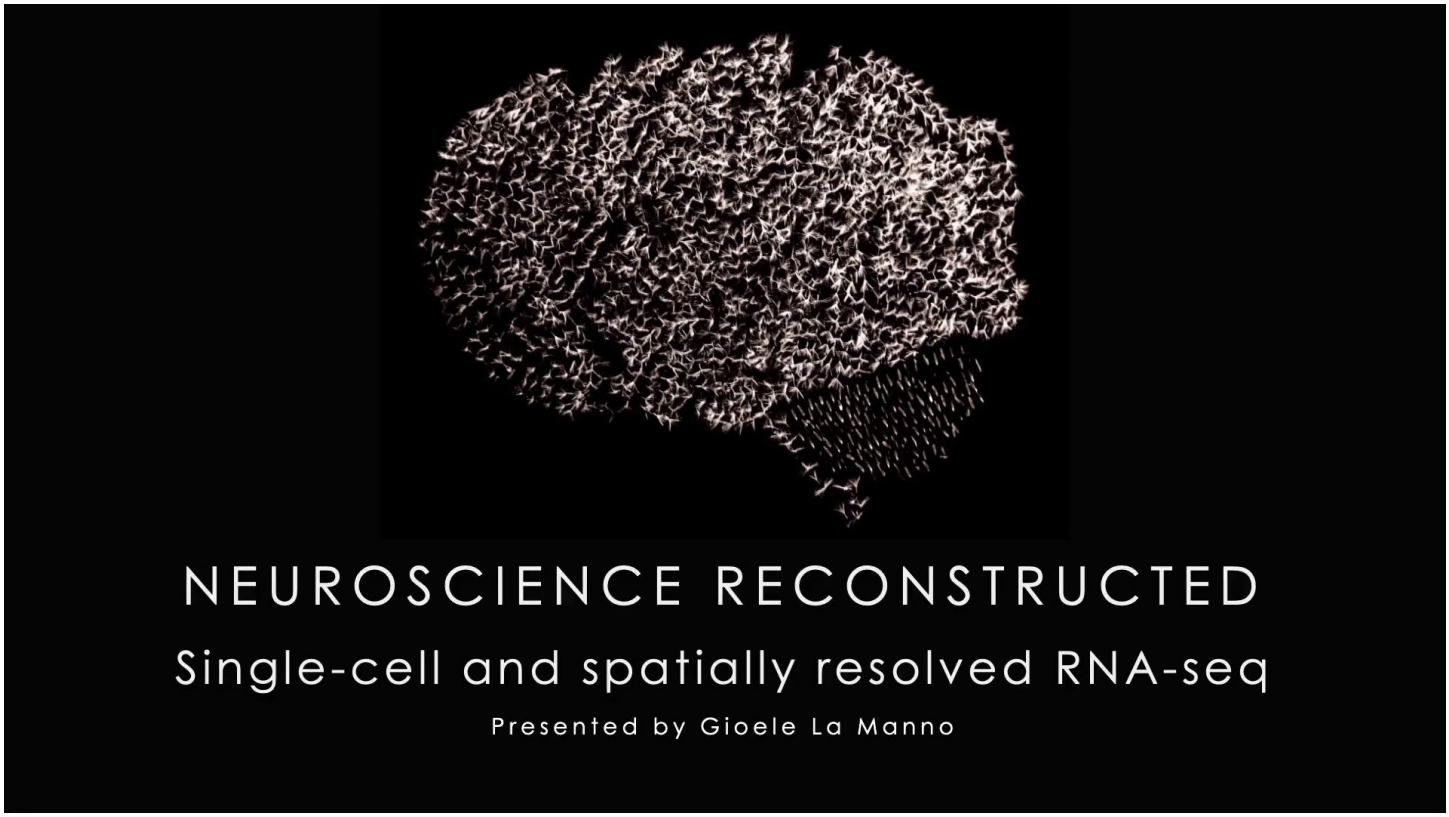
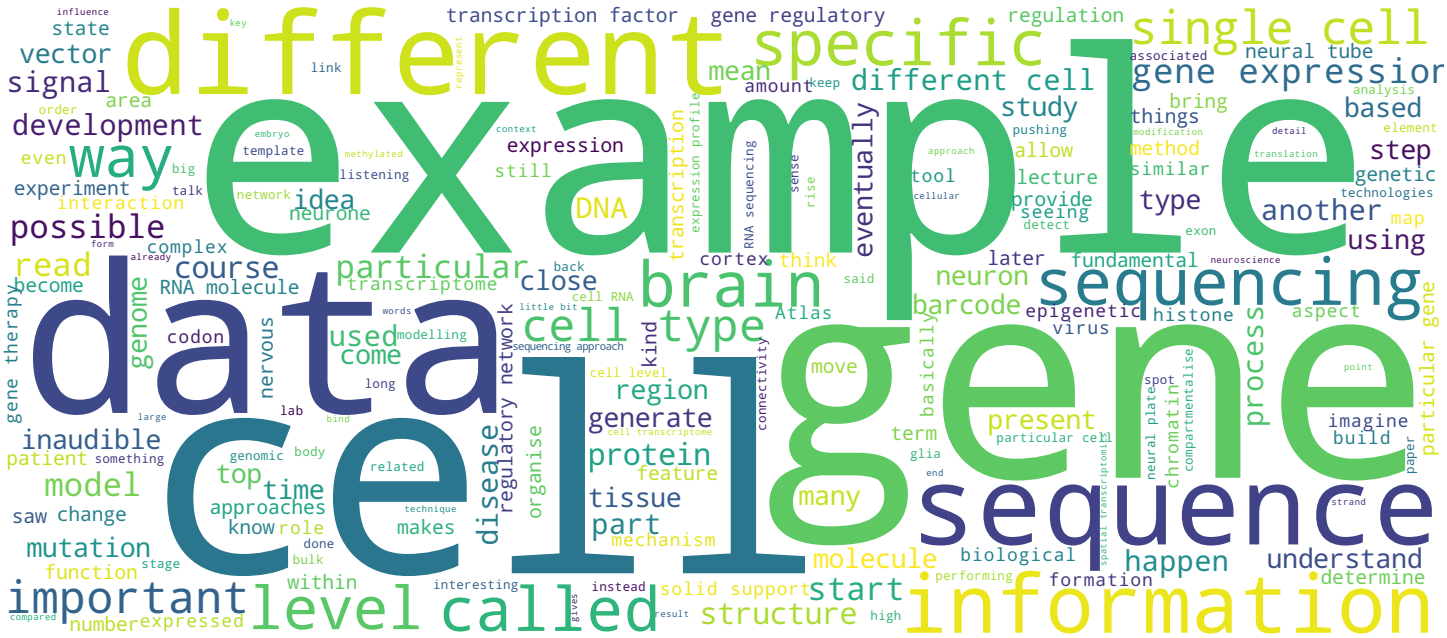
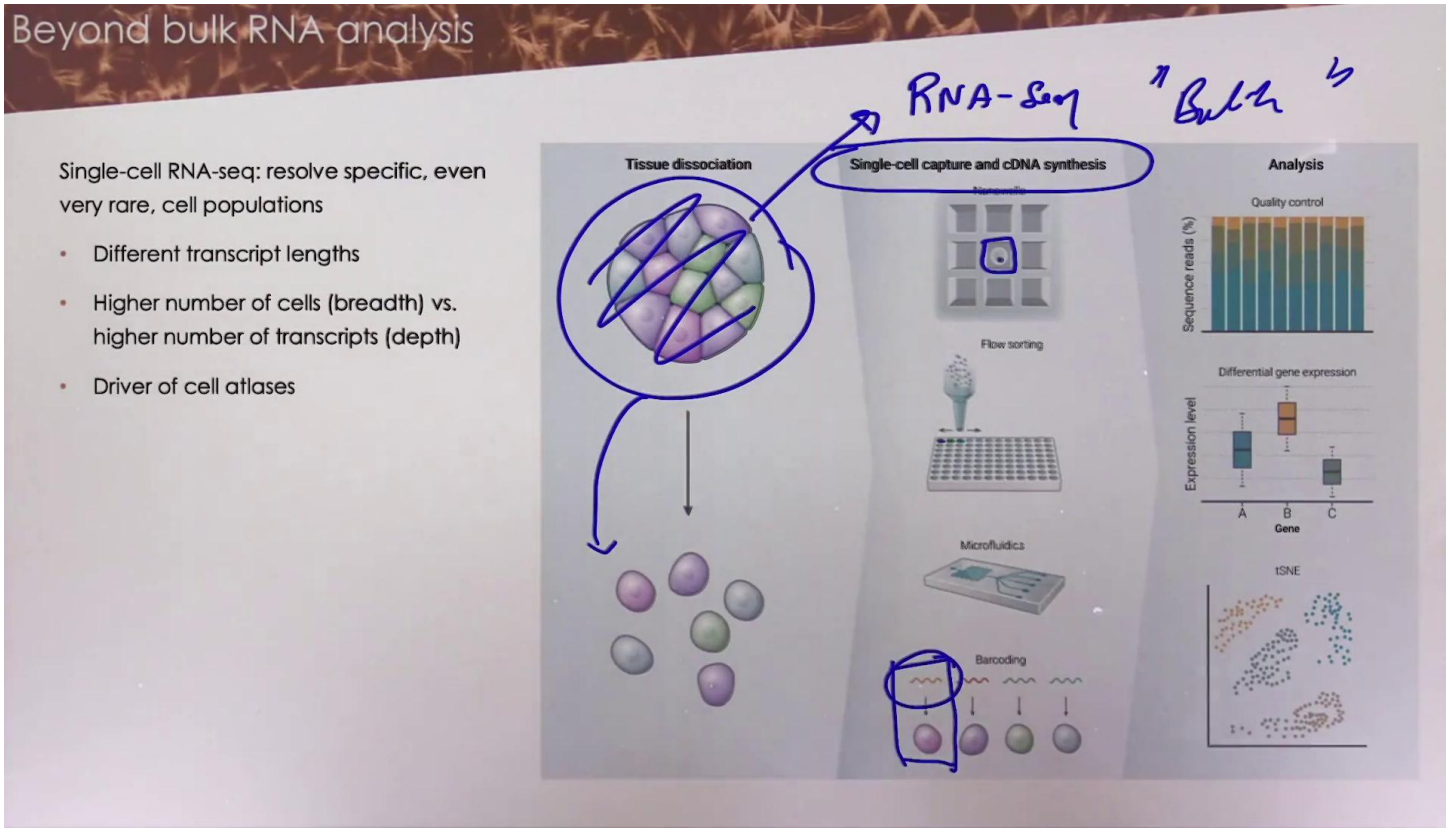



EPFL

**Search MOOC**

**Video**

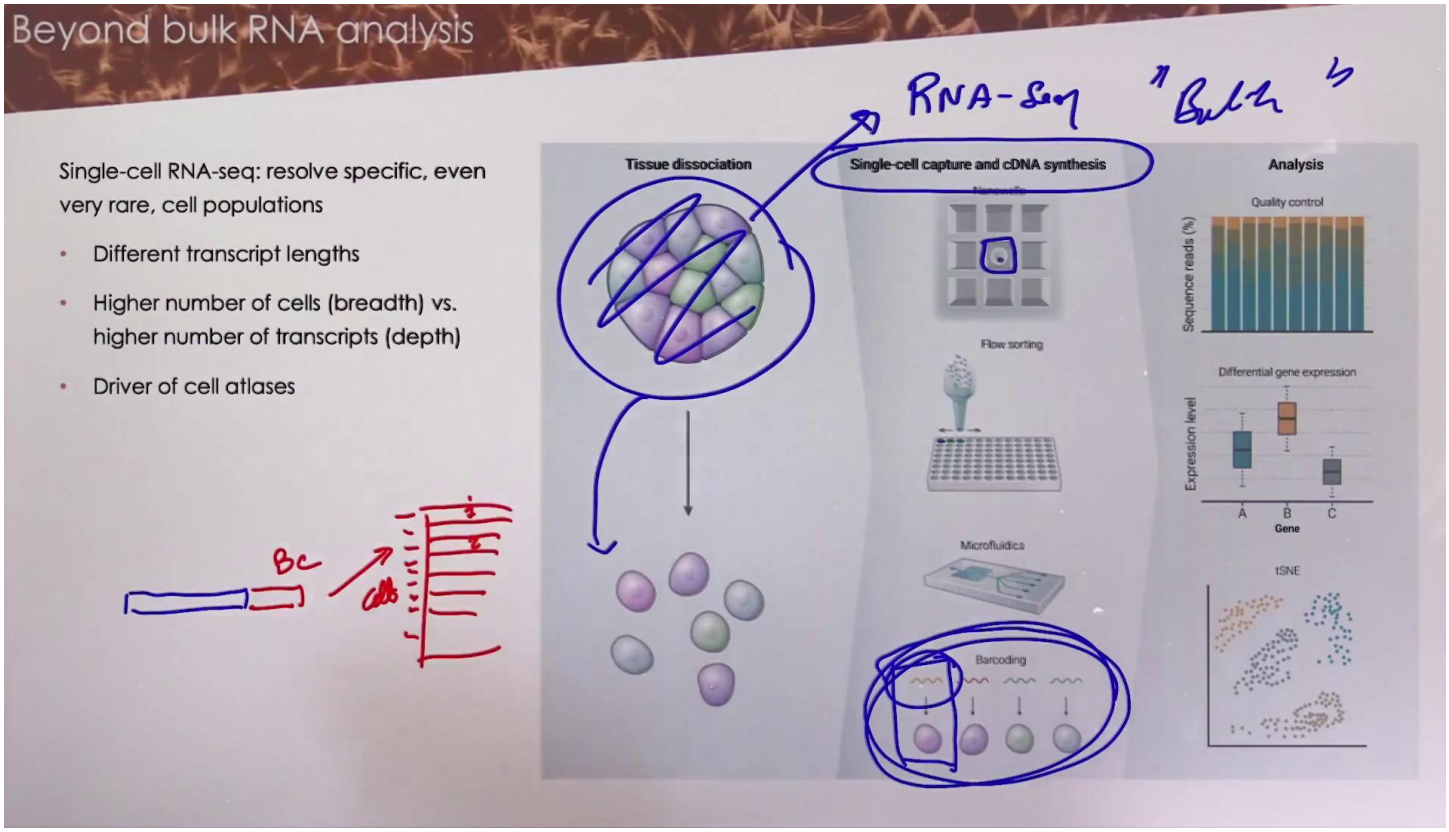



Over the last ten years, I can say we've been seeing a revolution of many of the genomics and techniques to move to the single cell level. Nowadays possible picking a tissue and instead of directly lighting it and sequencing a typical RNA-Seq what we call know this bulk sample. We can go about performing an actual single cell sequencing approach. This involves typically a first dissociation of a tissue in a single cell. Then, one among different strategies can be used to compartmentalise the cell, meaning that we need to have a system where each cell gets trapped in a particular compartment. Could be a nanowell, it could be through sorting into a multiwell plate, it could be to a microfluidics and to a particular microfluidics site or into a droplet. And once this cell is being isolated in a compartment, it's going to be possible to make the specific RNA of the cell addressable by barcoding it, in other words, going to be possible to provide to, for example, ligate or link by other means a particular oligonucleotide sequence to all the RNA molecules that are present in that particular cell.

Notes

Summary

0m 05s

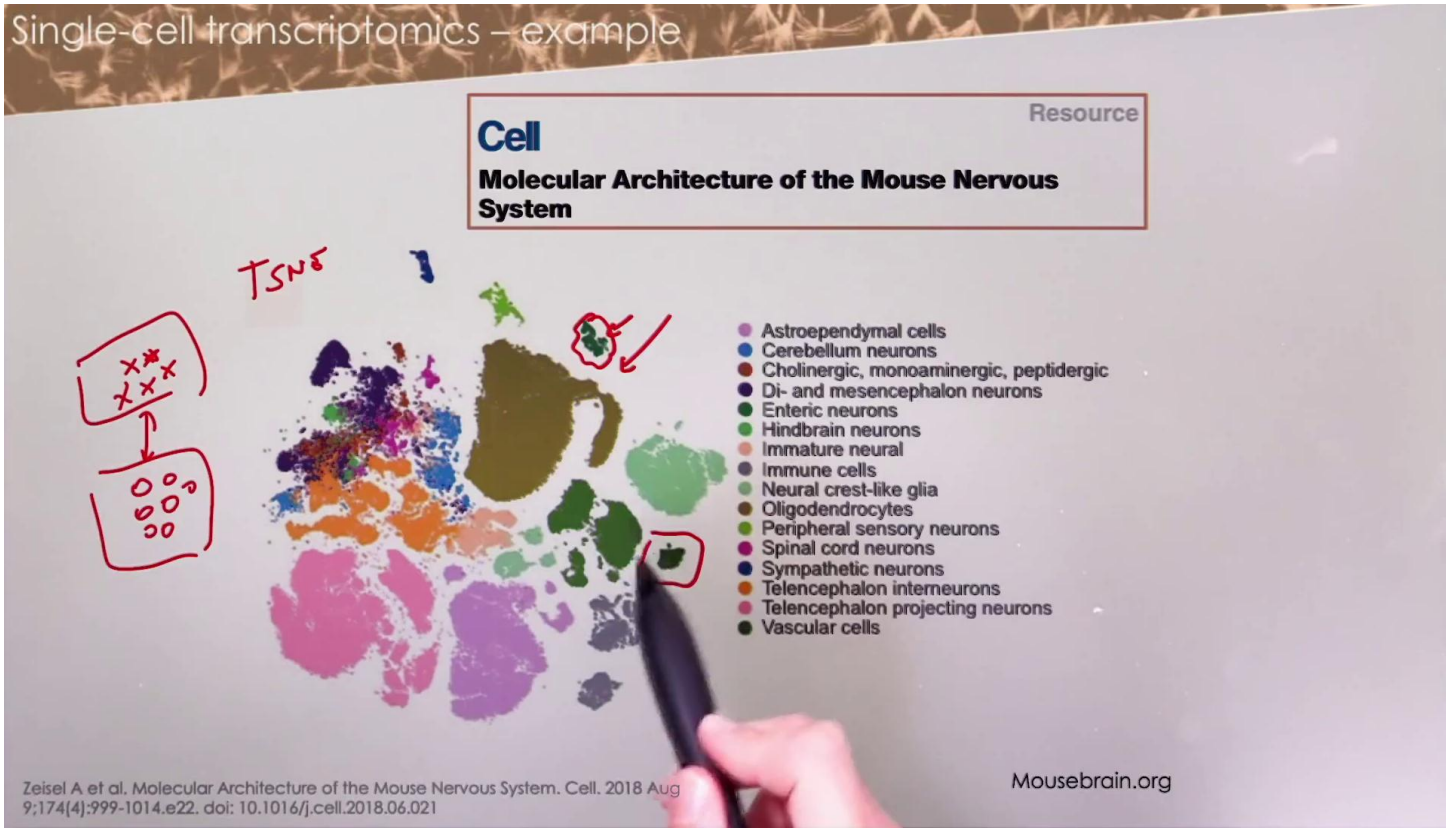



Now that every RNA molecule in the cell is stagged by a particular, a specific unique identifier for the cell, then possible to now bring all the content of the cells together, eventually they are marked by a specific identifier, and then go to sequencing and then at the sequencing level, it is going to be possible to do what is called demultiplexing, in other words, ask for each read that we sequence, what is the cell barcode and we can then put it in a different bucket. The idea is again, to sequence, we are going to sequence the body of the gene, but it's going to be associated to it… A specific barcode. And the barcode is going to tell us where we need to… to which cell the particular read belongs to, so that we can split, if you imagine a table, we can tally, we can basically count… the number of reads that relate to a gene that are expressed by one of the of many specific cells that we have been sequencing.

Notes

Summary

1m 35s

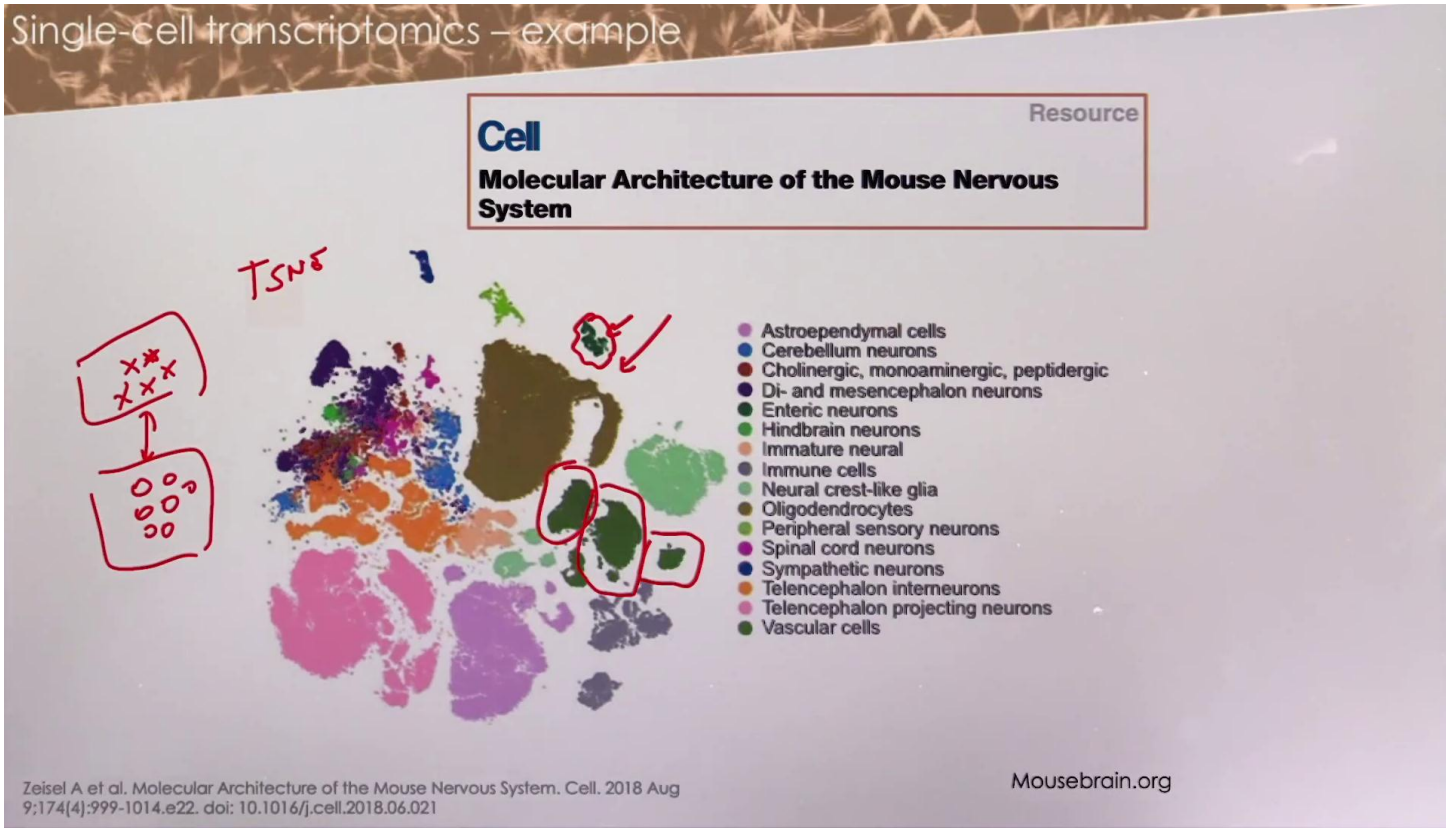



This is an example from the lab of Sten Linnarsson, that has been pioneering these technologies, to spearhead what is called cell type Atlas. The idea here, them showing a data set that is collected from sampling different areas of the mouse nervous system and performing of this different data single-cell RNA sequencing. So then we read the gene expression which said in each of these areas of the brain, and we can collect these into an Atlas that then we can… For example, displaying this layout where each dot that you see here in this dotty clouds represents a single cell transcriptome, and where the different cell transcriptomes are put next to each other by an algorithm called t-SNE, where if the gene expression profile for a particular cell is similar to the gene expression profile of another one, we represent them as [inaudible 00:03:53]. They are close to each other. What that means is that if their… cells are close to each other, going to be arrayed close to each other. And this other cell is different from those other guys, but close to this other cell. What we expect basically to happen if we array cells with this rule is that we're going to start to visualise cluster of cells that represent transcriptional different states, and therefore represent different cell types.

Notes

Summary

2m 46s

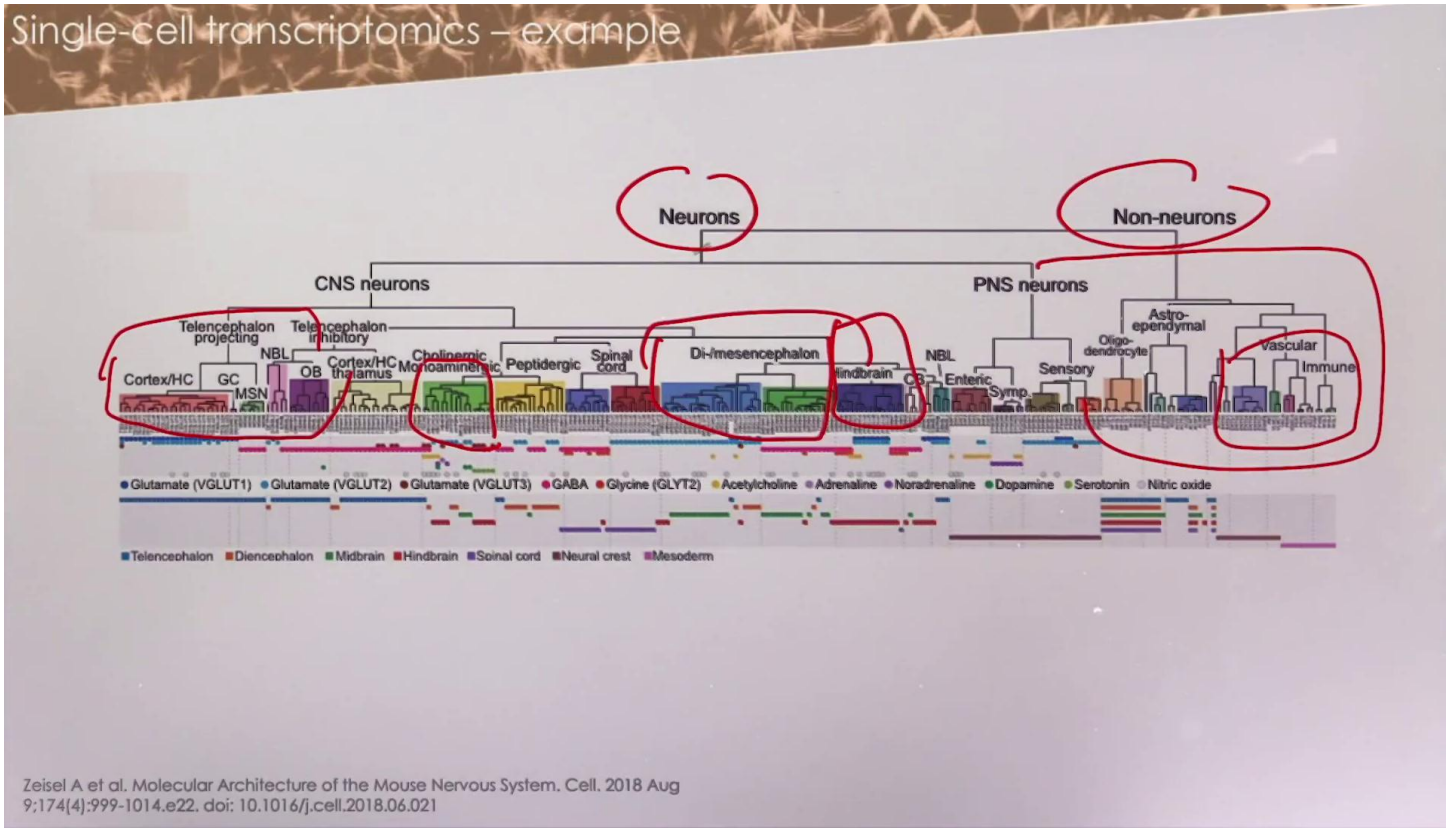



So in this typical visualisation, it's going to want just one of this step on the analysis of the single-cell transcriptomics, we can get information of what are the cell states or descend the cells types that are present to a tissue. And we can study them, we can basically profile them, build an Atlas or a map of the different cell types, and then study and identify what are the markers that makes them specific and makes them addressable, and go back to the lab and do a neurobiology on the specific different cell types.

Notes

Summary

4m 28s




So this is another representation of the result of such an approach. In this paper, for example, from Zeisel in the Linnarsson Lab, they've been defining up to 300 different cell population based on the transcriptome, then organise them into a taxonomy of the cell types. And you can see how the different cell type organised by, for example, the type of neurotransmitter that is being released. But at the larger scale they organise by their neurodevelopmental origin. You see here, for example, cell types that are specific to the telecephalon and some that are specific to the mesencephalon or to the hindbrain, and eventually there is a major split that you see at this very high level between the neurons in the glia and within the glia between different kind of [inaudible 00:05:56]. And eventually you come to the final group for other cells that are non-neuronal in origin, such as vascular and immune cells. You can get this very comprehensive… definition of the molecular heterogeneity that exist in a complex tissue such as the nervous system.

Notes

Summary

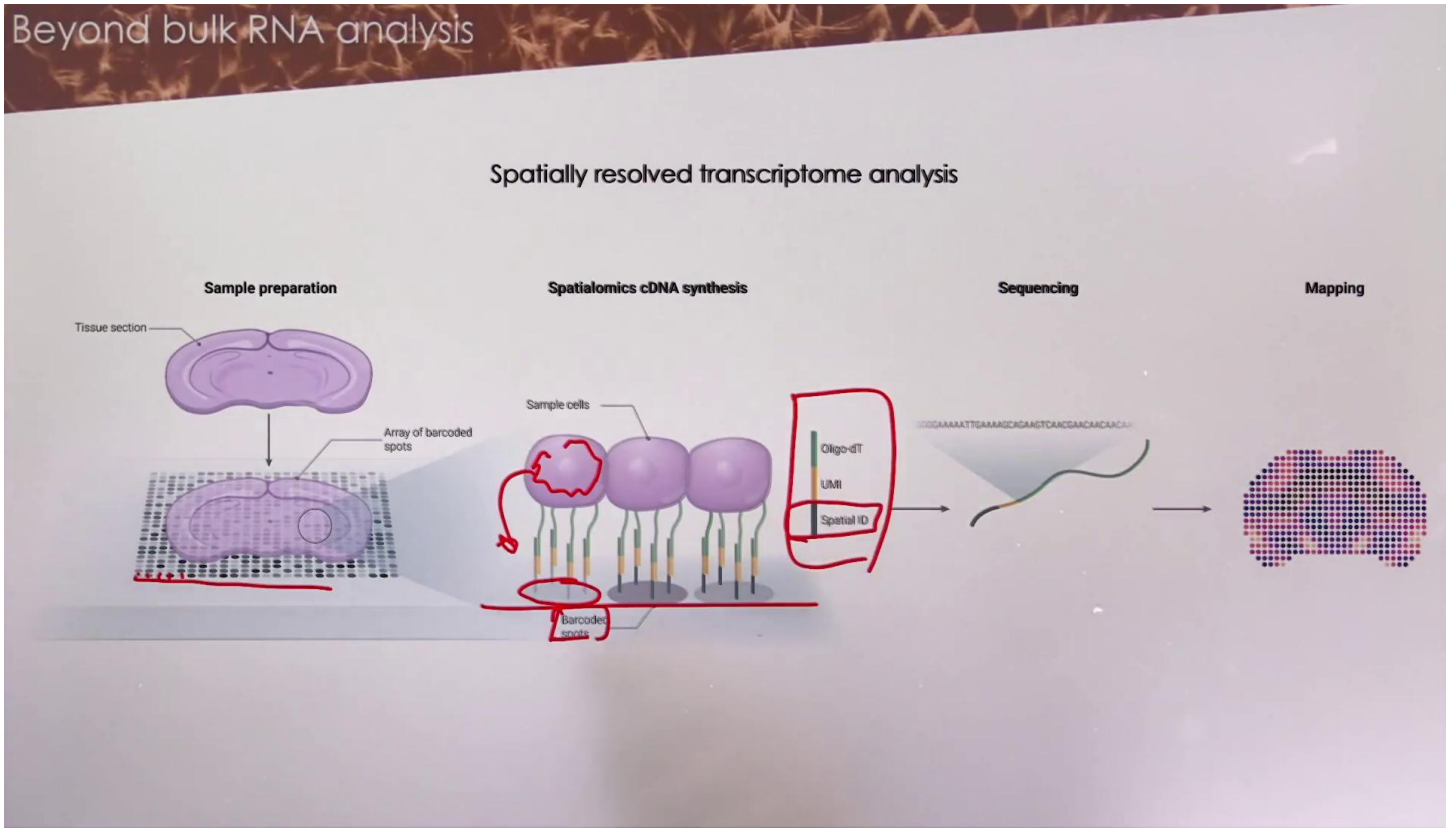
4m 59s

Beyond bulk and beyond even single-cell. The new frontier that is being explored a lot and has been exciting the community nowadays is what is generally referred as spatial transcriptomics. [inaudible 00:06:34] spatial transcriptomics says that… One of the aspects or the downside of the single-cell sequencing, whether it gives so much information and data originated in the single cell level, it loses because of technical reason. We saw this dissociate in tissue and compartmentalise, we lose information on the location of the cell. And this is important because if the logically is very relevant in development, with [inaudible 00:07:01], for example, is very spatially controlled, is important for things like connectivity to understand from which exactly area of the cortex, for example, the cell that we are sequencing the transcriptome of comes from. And to do this is not possible to retrieve that information with single-cell RNA sequencing, nor with bulk, so we need to use specifically grafted matter to retain that information.

Notes

Summary

6m 19s

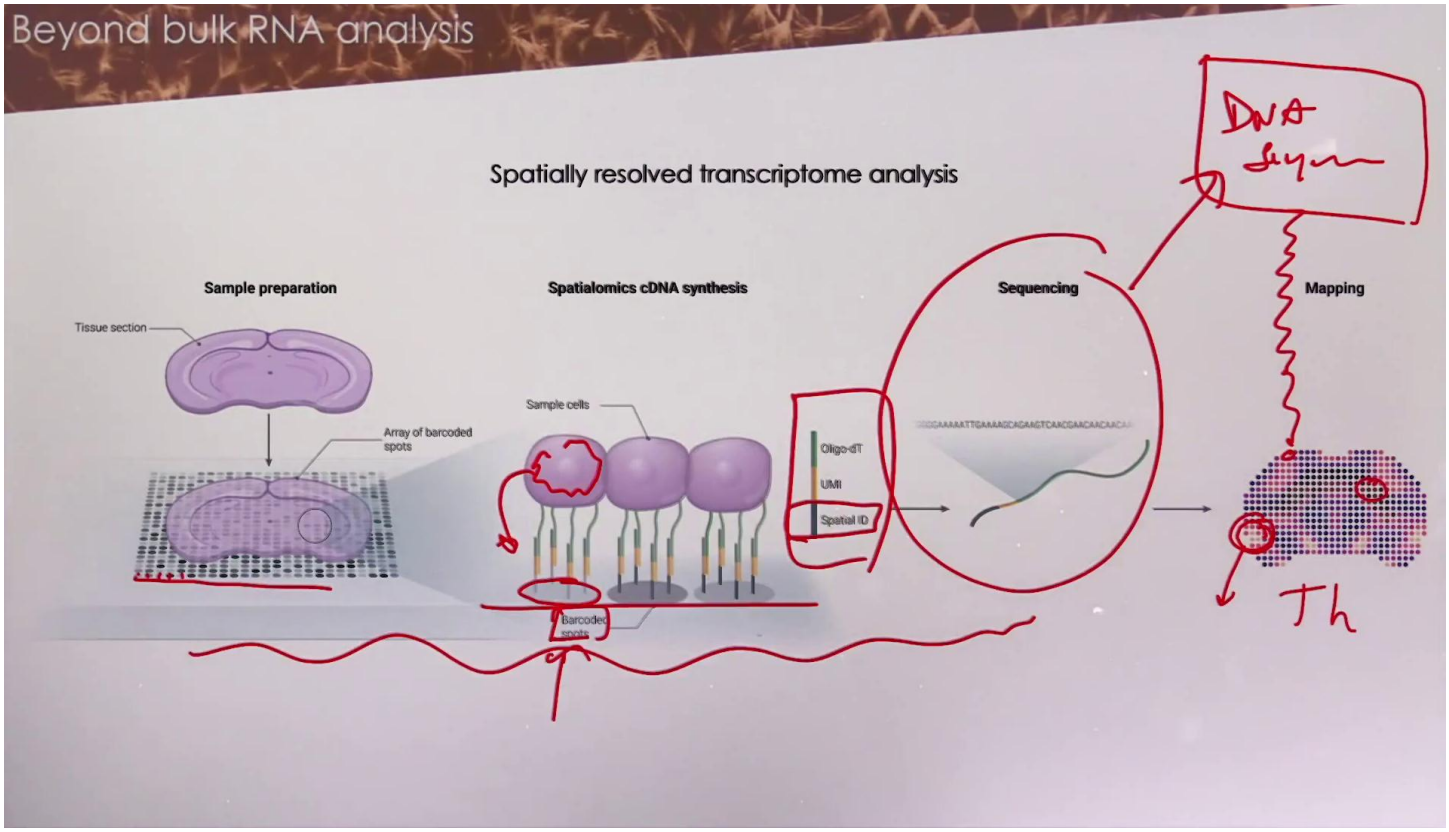



So in this setting, I will not give you… Cannot give you a complete summary of all the matter that just want to say that there exists different approaches, and one of the… approaches that is purely based on sequencing, but some approaches are actually not based on imaging, but some approaches that based on sequencing, for example, involve things like printing onto the micrometre sized feature where… In specific clusters on top of solid support, one can print nuclide with a specific barcoded sequence, so that each spot in this grid gets its own barcode. Then… One can engineer a way to transfer the RNA content into the cells. On top of this agreed by, for example, either by lysis and diffusion or electrophoresis, and one can then obtain reverse transcription on the solid surface. These eventually typically generate a library that has information on the gene, but it has also information on a spatial ID.

Notes

Summary

7m 27s

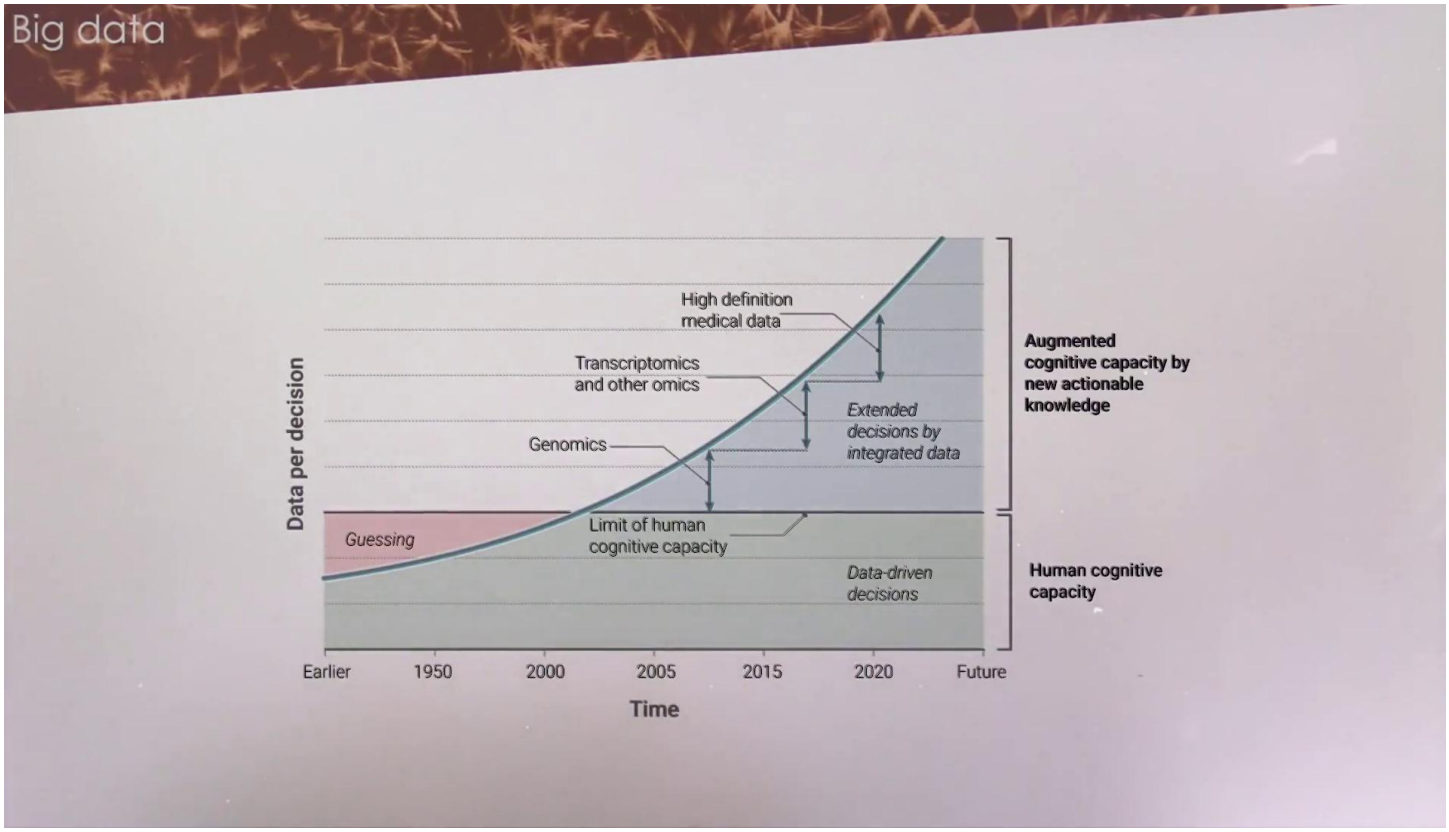



Through a set of steps involving the release of the cDNA from the solid support and the sequencing we can achieve then… Again, bring this pipeline to converge to a typical DNA sequencing approach that we have been seeing in the previous step, in the previous part of this lecture, and at the level of the bioinformatics analysis, we can ask to that particular sequence that we read, which gene does it map to, which particular barcode of sequence ID, localisation ID we find we read and therefore it is possible now to map back the information about the quantification of a particular gene expression onto a particular point in the green we can therefore say things like for a particular gene, say for example, tyrosine hydroxylase, we can detect more reads that come from barcode. The spots that correspond, for example, in this part of the cortex as compared with, for example, layer in the hippocampus. And we can do these kind of very spatially resolved and we can get this spatial resolved information for thousands of genes at the same time.

Notes

Summary

8m 51s

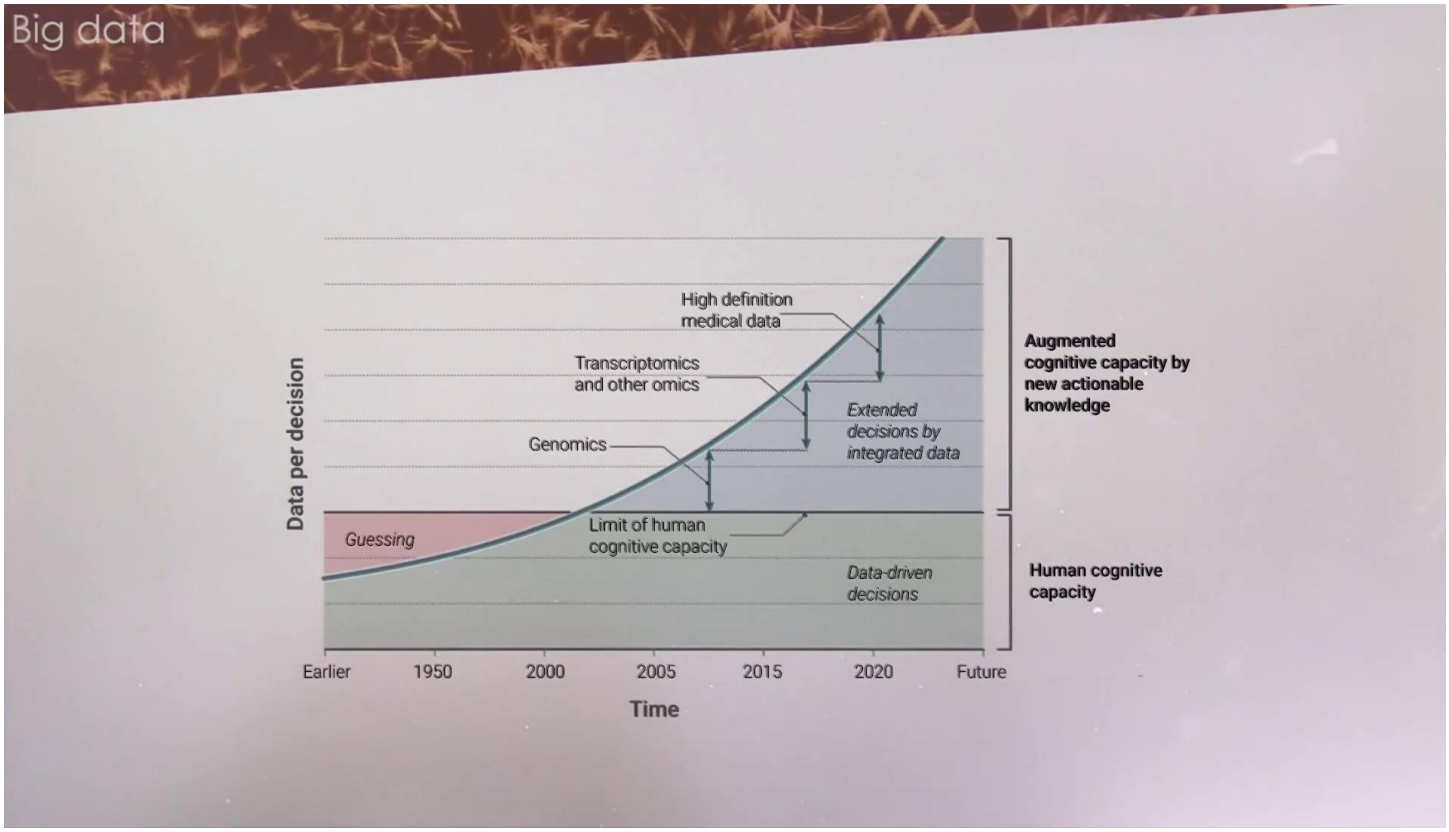



Eventually, and many of the proxies we've been discussing in this lecture, involve like pushing a crate, generating metal and pushing the technology to get more and more data. But with more data, we also start to lose full control from in terms of the limits of human cognitive capacity to handle this data directly and to manipulate it with… Do with common tools like we do with a spreadsheet. We need therefore to move to not only to programming base approaches, but also to way to analyse the data that is statistically found in the modelling… In models of these data, that can be statistical in nature, but they can also be more oriented to what we call machine learning. For example, pattern extraction and identification of specific correlation in specific mappings between features that we find in data, in our case, the level of expression of different genes, for example, or the amount to which a particular locus in the genome is methylated to… To explore and look at the covariance of different of such features and their association to, for example, an outcome variable or a covaiance such as disease versus control or the age of the sample.

Notes

Summary

10m 13s




And in this way… This is, I think, very important to keep in mind when tackling this kind of large size dataset, and they essentially encourage a whole new generation of biologists, neurobiology, neuroscientists, to tackle these problems in a very computationally driven and computationally oriented way to augment kind of the possibilities and information that we can extract from this data. And with this, I thank you for listening to this lecture, and I hope you are going to have a nice continuation with the exercise and for activities linked to this module of the course. Thank you for listening.

Notes

Summary

11m 49s