4th Talk at Café Eiles

Gene Over

Kaan Boztug, Jörg Menche,
Christoph Bock, Vanja Nagy

Gene Over

Investigate the genetic and epigenetic basis of human pathophysiological manifestations and phenotypes, from rare to common, from individuals to large cohorts, taking into account single-cell contribution, cell lineage, mosaicism and clonality and including network-based meta-analysis and disease modeling

Of how many people do you think you have analyzed some sort of genetic/genome information in the last ten years and how do you ensure data protection?

Bock: I wouldn't like to break down genomic impact into pure numbers, but in terms of our lab's focus, which has been on epigenetics over the last six years, samples have added up to something in the order of 10,000 – of which the vast majority is DNA methylation, but there is also a significant amount of chromatin accessibility data. This data helps us understand not only the genetic basis of a disease but oftentimes also provides insight into cancerous processes, which are also signs of immune dysregulation.

Menche: Is ensuring data protection a major issue in your group?

Bock: All genetic information and by extension epigenetic and transcriptional information is kept to very high standards of data protection. Nevertheless, it is important to recognize that data of this diverse nature can convey extensive private information, not only personal identity as used for forensic fingerprinting, but also DNA methylation information. You can accurately reconstruct for example smoking status and the number of smoking years from DNA methylation data. An American company is now even using DNA methylation-based age prediction to adjust life insurance premiums. So, this obviously means that we have to be very careful in terms of avoiding unauthorized access, while at the same time supporting international data exchange for research purposes as much as possible.

Menche: Let's turn to one of your best customers, I imagine that is Kaan's group. How many people have you analyzed over the last ten years?

Boztug: Just like Christoph, we don't care to define success so much by numbers because, as you know, our lab is majorly focused on solving individual patient cases, oftentimes theses are unique cases, rare disease patients that have a unique genetic defect. So even an n=1 can be the very significant start to a major discovery. That being said – just like Christoph's lab for epigenetics – we have certainly looked at a couple of thousand patients with different undiagnosed medical histories that we try to pinpoint down to a genetic etiology, mostly focusing on DNA sequencing to identify causative mutations for these patients.

Menche: And this trend is increasing, I imagine, if we wanted to pinpoint some sort of development.

Boztug: Absolutely. Our lab, like other labs, has seen a dramatic increase in the numbers of genetic samples that we are being referred to, which is because people are recognizing more and more how important this sort of deep genetic investigation is. And the other trend that we are seeing is that, obviously, with the costs of next-generation sequencing decreasing, people are moving more and more from panel sequencing towards exome sequencing, towards genome sequencing, because all of a sudden it is becoming affordable. So, genetic investigation is probing into increasingly deeper levels while the costs are decreasing on a per-base level.

Menche: What about data protection? The doctor of course has the highest obligations in this matter.

Getting a diagnosis has a major positive impact on the diagnostic odyssey of rare disease patients.

Boztug: Absolutely. Of course, protecting data is a very important endeavor and again, we are very happy to be part of a larger organization that has genetic and genomic medicine as one of its major hallmarks and thematic focuses. On the other hand, in my clinical practice, almost on a day-to-day basis, I experience that patients that are truly sick – or their parents in many cases because oftentimes small children are affected – are very interested in having people investigate them, even at the potential risk of genetic information being identifiable sometimes. In other words, there is an ethical debate that needs to consider the potential benefit of patients being included in a study, compared to the relatively low risk of information really being identified.

Boztug: I am, let's say, on the opposite end of the scale in terms of numbers. My team comes from a basic neurobiology point of view where we would like to take a particular gene and really study it in a lot of detail, especially using animal models. So, we spend a lot of time on each gene. That being said, we have started collecting and recruiting patients that are suffering from neurodevelopmental disorders and this is in the order of dozens and we hope to increase this to a much greater scale in the coming year. We have had some early level of success that is encouraging us to continue doing this. And of course, as everyone has mentioned, particularly because of the new data protection laws that have come out across Europe, we are incredibly aware and careful that all the data is protected appropriately and by European standards and the standards of the hospital that we are part of. I think Kaan touched upon this – especially in the rare disease community there is a lot of drive and there is a lot of push for data sharing and for certain information to be shared, not only between researchers themselves but within Europe and also outside of Europe. Because in order to confirm novel genes, you need to find patients that are suffering from similar clinical symptoms and we have to share these clinical details with other researchers and clinicians in order to confirm our findings and significantly shorten the diagnostic odyssey of each rare disease patient. And this poses a certain number of challenges – in order to find a balance to and preserve patient privacy, and for their data to be protected, to the best of our abilities, but at the same time not preventing us from actually being able to help them.

Menche: Much of the work we do in my group relies heavily on publicly available data, which Vanja also mentioned. If you consider, let's say, GWAS data that we have been using for many years and continue to use, I have used data on hundreds of thousands of patients. In a narrower sense, through actual collaborations with clinical partners and with physicians, this would be in the order of a few hundred at best, but it's also a trend that is increasing and certainly intensifying. What I can add on top of that is that we rely very much on the strict regulations of our collaboration partners that provide us with de-identified or anonymized data.

What has changed more recently? What is the limiting step in terms of analysis?

Menche: Clearly, at least according to my observations, the information that is being collected – I imagine also by our clinical cooperation partners – is becoming more and more high dimensional and diverse. So, there are more profiling expeditions where not only the genome is collected, but maybe also a proteome or transcriptome or metabolome. And a challenge related to this is, perhaps not so much computationally but really conceptually, how to integrate these diverse datasets that may span an entire range, from the molecular level to the cellular or organ level all the way to clinical phenotypes of entire cohorts. How to extract information and make something useful out of these data expeditions?

Boztug: Well, a number of things have changed. I think people are certainly becoming more aware that the multidimensional consideration of individual patient cases is what has the highest likelihood of succeeding in terms of pinpointing the genetic or even epigenetic etiology of a disease. In this regard, actually, our workflow has also changed. If I look back at the last couple of years, we are now trying much more to integrate not just genetics but also epigenetics, RNA sequencing, and a systems biology/network medicine perspective into our ways of characterizing patients, individual patients, but also diseases. The other trend that we see for immune diseases is that there is a greater need for medical doctors, alongside with researchers, to connect with different institutions, countries, the globe. It is still not unlikely that for some of the gene defects that we are working on in the Ludwig Boltzmann Institute for Rare and Undiagnosed Diseases, there may be patients in India, China and Japan that we just don't know about because there is no standardized database.

Nagy: So for neurological disorders, the challenge is obviously to functionally validate any novel genes that we find in a neurological disorder. Nervous tissue from humans required for genetic and molecular validation and research is not really accessible, so this is a major challenge in the field itself. And historically there has been a heavy reliance on animal models, but animal models have their limitations. Certain diseases, such as certain microcephalies for example, cannot be modelled in mice to the extent that would reflect the human condition. So, induced pluripotent stem cells have really taken over in laboratories. Transdifferentiated fibroblasts that are made into neurons and 3D organoid systems mimic early development of the brain. As a matter of fact, cerebral organoids, or mini-brains were developed right here in Vienna at IMBA and have created a very exciting boost and revolutionized the ways we can validate novel genes, understand neuropathologies and even search for patient-specific therapeutics. So this would certainly be something that is changing for the better in our field.

Bock: To me, 2018 has really been the year that artificial intelligence made it into the headlines, into common perception everywhere. That has tremendous potential, because obviously, artificial intelligence and machine learning, can consider extremely large datasets and make sense of them not only for prediction, but also increasingly in terms of understanding. But we are also hitting major roadblocks in this regard. Artificial intelligence experts are very difficult to get, we are educating far too few of them, so clearly there is the need for a lot more education, an outreach initiative to build a new generation of bioinformaticians, computer scientists, artificial intelligence technologists that can drive this change.

Any breakthroughs or new trends we can associate with 2018?

Bock: I think we have seen two dramatic breakthroughs in 2018. Perhaps the most visible – although I wouldn't want to call it a "breakthrough" – is the birth of the first CRISPR babies, the first gene-edited human beings. This was met with much deserved outrage in the scientific community, because in many ways it was done in direct violation of and contradiction to well-established standards enshrined in the scientific community's laws and regulations. Nevertheless, technologically, it was not entirely unexpected, because the methods have been there and available, so it was a matter of time before this was going to happen. I wish it had happened in a much more ethically sound and systematic, open, and regulated endeavor. CeMM has had a discussion on this, and issued a public statement, emphasizing not only the ethical limitations, but also the broader relevance of CRISPR technology. This was also picked up on by the Austrian Academy of Sciences. More generally, we are seeing a perhaps premature, but very clear push toward CRISPR genome editing technology with a real impact for human beings in general, and for medicine in particular. The other major breakthrough has been single-cell technology that is now everywhere. Single-cell sequencing but also increasingly high-resolution imaging technology coming together in initiatives like the Human Cell Atlas will provide great reference maps and a kind of starting point for dissecting diseases in full appreciation of their complexity.

Nagy: From a neuroscience perspective, to continue along the lines of what Christoph mentioned, the single-cell technologies that map out cell types in the brain are also a highlight for me. For instance the Allen Brain Atlas. They created a 3D model of total brain cell content, in incredibly comprehensive 3D imagery, available to researchers, and importantly, it is also updateable. This is really going to change the way the new and the young community of neuroscientists will be able to view the mouse brain. And eventually this will also happen for the human brain as well. On the other hand, to point out how little we still know about the brain, George Paxinos, who is one of the best-known cartographers of the human brain, or brain in general as we know it, has actually just discovered a new brain region or the connection between the spinal cord and the brain. So while we have all of this technology in front of us and we can do all these amazing things, we are still discovering different regions of the brain that can have huge implications. Particularly, this newly discovered region that is involved in motor processing and peripheral sensory input can have huge implications for diseases like Parkinson's.

Menche: I will also piggyback on something Christoph has already mentioned, which is artificial intelligence. For instance, Google's DeepMind team that came up with a program that taught itself to play chess, Go and shogi, and that can beat any human player. It has also found the first applications into medicine for certain specific applications where computer algorithms are now able to classify and diagnose – at least in laboratory experiments – at the level of an expert physician. Technically, most of these rely on huge amounts of imaging data, which is an area we are also exploring, for instance in terms of how to use imaging and classification and characterization of large-scale imaging experiments to improve our understanding of how drugs work.

Boztug: I am not sure if I can pinpoint one specific or two specific breakthroughs in 2018. In times where there is a lot of disagreement on political grounds, scientists have managed to actually collaborate quite well and maybe even better than before. This is true for people that work in biomedical research but also in more basic fields of research. And I think this is maybe another important message that we as scientists can send, namely regarding the significance of transnational and international collaboration for the betterment of humankind. This is not a breakthrough, but I think this is an important role that we as scientists have, particularly in times when some debates become particularly irrational in other fields, let's say, in politics.

Is there anything regarding this area that was achieved in your laboratory/at CeMM that you think is worth mentioning?

Bock: I think my two personal highlights of the last year were the following: first, we had a paper in *Nature Medicine* in the summer that described the disease progression in glioblastoma, the deadliest of brain tumors, based on a comprehensive Austrian cohort. This was the largest study that we have yet done in close collaboration with the Medical University of Vienna, involving several departments, including neuropathology, neuro-oncology, neurosurgery, radiology, etc. This study pulls together a massive dataset that combines not only epigenetic information but also brain imaging, digital pathology and several other layers into an integrative analysis. And it follows how these brain tumors first emerge and then, after surgery, chemotherapy, and radiation therapies, are initially cut back, but then reemerge as drug resistant tumors that ultimately kill the patient. The other point is that, over the last year, we have seen the wide adoption of our CRISPR single-cell sequencing technology, CROP-seq. Not only have we used it quite effectively in our lab, but it has also become a widely used technology, essentially as a new paradigm for CRISPR screening with many different labs already having published first results and showing the power of this technology that was originally developed at CeMM.

Boztug: Well, for us, 2018 has also been exciting in several regards. At the end of last year we were granted an ERC Consolidator Grant, which is important to motivate us to keep going and to keep our momentum and the focus on these rare gene defects which teach us so much about nature and the physiological functioning of molecules, cells and pathways and how they are important in a whole-body and in a disease perspective.

Menche: My team took part in two art exhibitions that I am very proud of, one here in Vienna and one in Munich. And we took our virtual reality holodeck on a road show to Ghent in Belgium, which was also really nice. That is definitely something I enjoy doing and that I would like to pursue further.

How important is it to be embedded in the MedUni/General Hospital setting?

Boztug: I cannot live in one world without the other, in practice. In reality, I am very happy both about the physical proximity of CeMM and the Medical University of Vienna and also the St. Anna Children's Hospital. I am really grateful that I have actually seen open doors in both directions from the research world to the clinics and vice versa. I think when we do research, we have an obligation, especially as medical doctors and physician scientists, to do something meaningful for humankind and for sick patients. And we are trying to use this opportunity to the best of our abilities for the betterment of these sick patients. So yes, I think it is crucially important for an institution that focuses on molecular medicine with a relevance to diseases to be embedded in such a setting and I am really happy about this specific setting that we have entered into. That is very fruitful and that is why we are here.

Bock: The close collaboration with groups and clinics of the Medical University of Vienna and the General Hospital is a major driving force in our research. I have already mentioned the collaboration on glioblastomas. We also have very strong collaboration in the area of hematology, studying leukemias, with several papers in the pipeline. And beyond that, we are working closely with the St. Anna Children's Hospital and the Children's Cancer Institute which gives rise to exciting endeavors combining clinical cohorts and clinical insights with high-throughput technologies in pediatric cancer research. Through the Biomedical Sequencing Facility of CeMM and the Medical University of Vienna, our technology platform for next-generation sequencing, we have also supported a large amount of collaborations, interactions and individual projects with sequencing services and have increasingly contributed to making the vision of "whole genome medicine" a reality in many fields of clinical and experimental medicine.

Nagy: It is absolutely critical that research labs involved in basic science are in direct contact and in the direct vicinity of the clinics, because we can provide the kind of assays that would be able to dig into patient sequencing databases and to prioritize genes that can aid in some of the diagnostic and therapeutic decisions that clinicians can make. Our work together with systems biologists for example, can utilize known gene networks in any particular group of disorders. I think psychiatric disorders are a perfect example and can even help redefine the disease. We are beginning to understand that clinical findings are simply not enough. Cognitive symptoms can be very tricky, and it is necessary to actually find the biological underlying cause of the disease in question. So it is absolutely critical that labs like ours that are involved in basic science are actually embedded within the clinical community.

Menche: In my own team, many people do not have a very strong biomedical background. So we, have physicists, mathematicians, and some people with a more chemical background or an artist, and to us it is absolutely critical to be immersed in this environment, in order to keep on track in our very interdisciplinary endeavors. To ensure that we ask the right questions.

By when do you think will we see impact of this research in medical practice?

Bock: In terms of therapy development, the path from target discovery to approved medicine has been traditionally very long. 10 years is often given as a realistic estimate. However, we do see a whole new generation of academic labs that are now involved in the development of new small-molecule drugs. There are other treatments, too, which can be developed quite efficiently in a university setting. CRISPR gene therapy is one, cell-based therapy such as CAR T cell therapies another, mRNA-based therapies might be yet another, or antisense oligo nucleotide-based therapies.

Menche: The research part in my group that is by far closest to medical practice, is done together with Kaan and Vanja. One thing that I am dreaming about or that I am being delusional about is that, perhaps within the next 5 years, I would love to see some of our virtual reality in a clinical research setting. Kaan and Vanja, when is it that our methods are going to be useful in that sense?

Nagy: I was really just going to say for a rare disease patient, whose case is solved, the impact is right now. Getting a diagnosis has a major positive impact on the diagnostic odyssey of rare disease patients. We try to, intelligently, with your help, pick the certain genes that we would screen for, that we would functionally validate, and then go back to the undiagnosed patient files of our clinical partners and see if we can help them identify causative genes already prioritized by standard sequencing techniques.

Boztug: I fully agree. We have seen tremendous impact. The mere fact of having a diagnosis often changes a lot for patients in terms of their psychological well-being. And in some selected cases, this has a direct implication for treatment, because we know some treatments that are efficient for certain diseases. Or actually, we have showcased that in some cases, we can repurpose existing drugs, in a sort of precision medicine type of approach, and use them for specific individual patients given their specific genetic background. That being said, I think the major challenge in the field, for us and also for many other people, remains to be more efficient, to really make that step into targeted therapies for these patients. This will be a major effort that will require a combined effort of treating physicians, researchers and chemical biologists, but that also needs support and collaboration with the pharma industry. All together in an alliance to best serve patients and their individual needs.