Genomics is becoming the next significant challenge for privacy [10]. The price of a complete genome profile has plummeted below $100 for genome-wide genotyping (i.e., the characterization of about one million common genetic variants), which is offered by a number of companies. This low cost of DNA sequencing will break the physician/patient connection and it can open the door to all kinds of abuse, not yet fully understood.

The spectacular drop in the cost of genome sequencing is also paving the way to personalized health. Diagnoses will improve, treatments will be more appropriately applied, and medicine will be increasingly preventive. However, the impact on privacy and security is unprecedented, because (i) genetic data can be used to identify individuals (ii) genetic diseases can be unveiled, (iii) the propensity to develop specific diseases (such as Alzheimer’s) can be revealed, (iv) a volunteer accepting de facto to have his genome made public (already the case) can leak substantial information about their ethnic heritage and the genomic data of their relatives (possibly against their will), and (v) complex privacy issues can arise if DNA analysis is used for criminal investigations and insurance purposes. Such issues could lead to genetic discrimination (e.g., ancestry discrimination or discrimination due to geographic mapping of people). Even though legislation such as the EU General Data Protection Regulation (GDPR) and the US Genetic Information Non-discrimination Act (GINA) aim at mitigating the risks of such discrimination, it is clear that genomes also need to be protected by technical means.

At EPFL, we have been working on different aspects of genome privacy in strong collaborations with genomic researchers and medical doctors. In particular, we focus on the following main research directions:

- Data protection: we make use of decentralized cryptographic protocols to prevent data leakage [1,2]
- Privacy: we quantify the risk of inference attacks and propose techniques to minimize it.
- Protecting and Evaluating Genome Privacy in Medical Tests and Personalized Medicine [1,2,3,4,7,9,13,14,17]
- Privacy-Preserving Processing of Raw Genomic Data [6,16]
- Quantification and Protection of Kin Genomic Privacy [5,8,10,14,15].

Academic Partners: University Hospital in Lausanne (CHUV), Vanderbilt University, Cornell Tech, Stanford University, Case Western University, Swiss Data Science Center, Scripps Institute, Indiana University at Bloomington.

Contacts: Prof. Jean-Pierre Hubaux jean-pierre.hubaux@epfl.ch


