

PROTECTING DATA FOR PERSONALIZED HEALTH AT EPFL, LAUSANNE

<https://lca1.epfl.ch/pdph/>



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

- [1] J.L. Raisaro, J.R. Troncoso Pastoriza, M. Misbach, J. Sa Sousa, S. Pradervand, E. Missiaglia, O. Michielin, B. Ford and J.-P. Hubaux, "MedCo: Enabling Privacy-Conscious Exploration of Distributed Clinical and Genomic Data", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2018.
- [2] J.L. Raisaro, G. Choi, S. Pradervand, R. Colsenet, N. Jacquemont, N. Rosat, V. Mooser, and J.-P. Hubaux, "Protecting Privacy and Security of Genomic Data in i2b2", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2018.
- [3] Z. Huang, J. Fellay, Z. Kutalik and J.-P. Hubaux, "SQC: Secure Quality Control for Meta-Analysis of Genome-Wide Association Studies", To appear in Bioinformatics.
- [4] J. Sa Sousa, C. Lefebvre, Z. Huang, J.L. Raisaro, C. Aguilar, M.-O. Killijian, J.-P. Hubaux, "Efficient and Secure Outsourcing of Genomic Data Storage", To appear in BMC Medical Genomics.
- [5] J.L. Raisaro, F. Tramer, et al. "Addressing Beacon re-identification attacks: quantification and mitigation of privacy risks", Journal of the American Medical Informatics Association, Feb. 2017.
- [6] Z. Huang, E. Ayday, H. Lin, R. S. Aiyar, A. Molyneaux, Z. Xu, J. Fellay, L. M. Steinmetz, and J.-P. Hubaux, "A Privacy-Preserving Solution for Compressed Storage and Selective Retrieval of Genomic Data", Genome Research. 2016 Dec 1;26(12):1687-96.
- [7] P.-J. McLaren, J.L. Raisaro, et al., "Privacy-Preserving Genomic Testing in the Clinic - A Model Using HIV Treatment", Genetics in Medicine. Jan. 2016. Nature Publishing Group.
- [8] F. Tramèr, Z. Huang, E. Ayday, and J.-P. Hubaux, "Differential Privacy with Bounded Priors: Reconciling Utility and Privacy in Genome-Wide Association Studies", accepted by 22nd ACM Conference on Computer and Communications Security (CCS), Denver, CO, USA, October 2015.
- [9] Z. Huang, E. Ayday, J. Fellay, J.-P. Hubaux, and A. Juels, "GenoGuard: Protecting Genomic Data against Brute-Force Attacks", 36th IEEE Symposium on Security and Privacy (S&P), San Jose, CA, USA, May 2015.
- [10] M. Humbert, E. Ayday, J.-P. Hubaux, and A. Telenti, "On Non-cooperative Genomic Privacy", 19th International Conference on Financial Cryptography and Data Security (FC), 2015.
- [11] E. Ayday, E. D. Cristofaro, J.-P. Hubaux, and G. Tsudik, "Whole Genome Sequencing: Revolutionary Medicine or Privacy Nightmare?", IEEE Computer, vol.48, no. 2, pp. 58-66, February 2015.
- [12] M. Naveed, E. Ayday, E. W. Clayton, J. Fellay, C. A. Gunter, J.-P. Hubaux, B. A. Malin, and X. Wang, "Privacy in the Genomic Era", ACM Computing Surveys, vol. 48, num. 1, 2015.

- [13] M. Naveed, S. Agrawal, M. Prabhakaran, X. Wang, E. Ayday, J.-P. Hubaux and C.A. Gunter, "Controlled Functional Encryption," In proceedings of 21st ACM Conference on Computer and Communications Security (CCS), Scottsdale, AZ, USA, November 2014.
- [14] M. Humbert, E. Ayday, A. Telenti, and J.-P. Hubaux, "Reconciling Utility with Privacy in Genomics," In Proceedings of ACM Workshop on Privacy in the Electronic Society (WPES), Scottsdale, AZ, USA, November 2014.
- [15] M. Humbert, E. Ayday, A. Telenti, and J.-P. Hubaux, "Addressing the Concerns of the Lacks Family: Quantification of Kin Genomic Privacy," In Proceedings of 20th ACM Conference on Computer and Communications Security (CCS), Berlin, Germany, November 2013.
- [16] E. Ayday, J. L. Raisaro, U. Hengartner, A. Molyneaux, and J.-P. Hubaux, "Privacy-Preserving Processing of Raw Genomic Data," In Proceedings of 8th Data Privacy Management (DPM 2013) International Workshop (in conjunction with ESORICS), Egham, UK, September 2013.
- [17] E. Ayday, J. L. Raisaro, P.J. McLaren, J. Fellay, and J.-P. Hubaux, "Privacy-Preserving Computation of Disease Risk by Using Genomic, Clinical, and Environmental Data," In Proc. of USENIX Security Workshop on Health Information Technologies (HealthTech '13), Washington, DC, USA, August 2013.