PRAXIS

Schweizerische Rundschau für Medizin
Revue suisse de la médecine

10

7. Mai 2014 · 103. Jahrgang
le 7 mai 2014 · 103e année

www.praxis.ch
ISSN 1661-8157

PRAXIS ist gelistet in MEDLINE, EMBASE und Scopus

Therapie der Anaphylaxie

Schwerpunkt: Genomik/Génomique

Genomik und personalisierte Medizin

Enjeux éthiques et sociaux de la médecine génomique

Patient Privacy in the Genomic Era

Médecine génomique et Maladies infectieuses

Médecine génomique et oncologie

Polyurie-Diagnostik

«Rational Testing» in Zusammenarbeit mit BMJ

Zertifizierte Fortbildung CME

In Zusammenarbeit mit: BMJ, Cochrane Schweiz, The Cochrane Collaboration

Editorial

Continuing Medical Education

Mini-Reviews: Genomik/Génomique

Verlag
Verlag Hans Huber
Hogrefe AG
Länggass-Strasse 76
Postfach, 3000 Bern 9
Tel. 031 300 45 00
Fax 031 300 45 90
www.verlag-hanshuber.com
www.praxis.ch

Leitende Redaktorin
Valérie Herzog
Tel. 031 300 45 76
Fax 031 300 46 27
redaktion@praxis.ch

Herstellung
Karolina Andonovska
Tel. 031 300 45 75
Fax 031 300 46 27
praxis@hanshuber.com

Anzeigenleitung
Brigitte Niederberger
Tel. 031 300 45 69
Fax 031 300 45 91
inserate@hanshuber.com

Abonnemente
Tel. 031 300 45 55
Fax 031 300 45 91
zeitschriften@hanshuber.com

Satz und Druck
AZ Druck und
Datentechnik GmbH
Heisinger Strasse 16
87437 Kempten (Allgäu)
Deutschland

Abonnementspreise
(inkl. Porto und Versand)

Private	CHF 218.–
Assistenzärzte	CHF 113.–
Studenten	CHF 102.–
Institute	CHF 411.–
Einzelheft	CHF 30.90

+ Porto und Versandgebühren

Erscheinungsweise
14-täglich, jeweils mittwochs

PRAXIS ist gelistet in MEDLINE, EMBASE und Scopus.

ISSN-L 1661-8157
ISSN 1661-8157 (Print)
ISSN 1661-8165 (online)

PRAXIS

School of Computer and Communication Sciences, Laboratory for Communications and Applications (LCA1), EPFL Lausanne

Jean Louis Raisaro, Erman Ayday, Jean-Pierre Hubaux

Patient Privacy in the Genomic Era

Datenschutz in der Genomik-Ära

Abstract

According to many scientists and clinicians, genomics is taking on a key role in the field of medicine. Impressive advances in genome sequencing have opened the way to a variety of revolutionary applications in modern healthcare. In particular, the increasing understanding of the human genome, and of its relation to diseases and response to treatments brings promise of improvements in better preventive and personalized medicine. However, this progress raises important privacy and ethical concerns that need to be addressed. Indeed, each genome is the ultimate identifier of its owner and, due to its nature, it contains highly personal and privacy-sensitive data. In this article, after summarizing recent advances in genomics, we discuss some important privacy issues associated with human genomic information and methods put in place to address them.
Key words: genome – privacy – information security – cryptography – anonymization

Introduction

Privacy issues have a rather long history. The photo camera, introduced at the end of the 19th century, was the first revolutionary observation and identification tool that threatened the privacy of an individual. Since then, several other tools have become widespread, including: video cameras, credit cards, Web browsers, and mobile phones. These tools reveal our presence and habits in various spheres of life, as well as our communication and mobility patterns [1]. DNA sequencing greatly exacerbates this problem, as the genome represents our ultimate biological identity. By combining genomic data with information about a person's environment or lifestyle, it is possible (to some extent) to infer that individual's phenotype.

The genomic era began in April 2003, when the Human Genome Project was declared complete. Since then, as a result of the impressive decrease in genome-sequencing costs and the rapid development of new-generation sequencing technologies, medicine has been undergoing a genomic revolution. It is not unrealistic to believe that, in a near future, most individuals will have their genomes sequenced. Thus, they will be able to benefit from preventive diagnoses and treatments tailored to their genetic makeup. Even though the understanding of the complex relation between genome and health is still superficial and much more progress has to be made in biomedical research, it is now possible to collect, store, process and share genomic data in a way that was unthinkable a decade ago. However, the rise in availability, use, and sharing of such information raises many serious ethical and privacy concerns.

In general, access to genomic data prompts some important privacy concerns: (i) DNA reflects information about genetic conditions and predispositions to specific diseases such as Alzheimer's, cancer, or schizophrenia, (ii) DNA contains information about ancestors, siblings, and progeny, (iii) DNA (almost) does not change over time, hence revoking or replacing it (as with other forms of identification) is impossible, and (iv) DNA analysis is already being used both in law enforcement and healthcare, thus prompting numerous ethical issues. Furthermore, it is hard to assess or estimate the extent of the personal information that could be extracted or derived from the genome in the future.

Traditional approaches to privacy, such as de-identification or aggregation [2], are ineffective in the genomic context because the genome itself is the ultimate identifier [3]. For instance, a recent study by Gymrek et al. [4], published in *Science*, demonstrates the feasibility of re-identifying DNA donors from a public research database by using information available from popular genealogy Websites.

In this article, we first focus on the current uses of genomic data and the possible privacy threats. We then describe some of the techniques recently designed to protect the genomic data.

Abbreviations used in the article:

DNA	Deoxyribonucleic acid
DTC	Direct-to-Consumer
EHR	Electronic Health Record
OSN	Online Social Network
PET	Privacy-Enhancing Technology
SPU	Storage and Processing Unit
WGS	Whole Genome Sequencing

DOI 10.1024/1661-8157/a001657

Uses of genomic data and potential threats

The DNA sequence of an individual is composed by over 3 billion base pairs distributed along 23 pairs of chromosomes. Most of the DNA sequence is conserved across the whole human population and it is estimated that no more than 0,5% of the DNA differs between any two individuals [5]. Yet, it is these differences that influence an individual's health status and other aspects. In recent years, the use of genomic data has increased dramatically. Its applications can be grouped in four main areas: healthcare, genomic research, direct-to-consumer services, and legal and forensic genomics.

Healthcare

It has been proved that mutations in an individual's genomic makeup can influence his or her health status. In particular, genetic variations can be associated with change in the susceptibility to a certain disease and in the response to a pharmaceutical agent. Thus, early diagnosis and treatment can be tailored to an individual's genetic makeup [6]. Examples of Swiss initiatives in the digitization process of health records are monDossierMedical.ch [7] and healthbank.ch [8].

Genomic Research

On a weekly basis, researchers discover new associations between the genome and a number of disorders and variable responses to treatments. The dramatic decrease in the cost of genome sequencing has made it increasingly possible to collect, store, and computationally analyze genomic sequencing data on a fine-grained level, as well as over populations on the order of millions [9].

Direct-to-Consumer Services (DTC)

Until recently, genome sequencing was a complex and expensive process, hence a prerogative of only large research laboratories or diagnostic centers. But, in the past few years, there has been a rise in DTC services for medical data in general, and genomic data in particular. These services have made it affordable for individuals to become directly involved in the collection, processing, and even analysis of their medical and genomic data. Some well-known examples are 23andMe [10] and Counsyl [11] that provide their customers with reports, based on their genotype, about the risk of developing certain diseases.

Legal and Forensic Genomics

Given that genomic data has a static nature and uniquely identifies an individual, this information is used also for investigative purposes.

Although the widespread availability and use of genomic information is an outstanding opportunity for medicine, the impact on privacy is unprecedented [1]. Of course, tight legislation regulates the activities of companies and hospitals that perform DNA sequencing and store genomic data, but it is difficult to protect them against the misdeeds of a hacker or a disgruntled employee.
The leakage of genomic information can pave the way to a variety of abuses and threats. For example, health insurance companies could obtain the genetic information of their customers and deny their services to people with a high susceptibility of developing a chronic disease, or employers could hire applicants based on their genetic features. Access to this information could engender genetic discrimination as depicted by the popular sci-fi movie of 1997 «GATTACA». In addition, a medical institution aiming at setting up a medical study on genomic data could be severely discredited if participants' genomic information were leaked or compromised.
Furthermore, integration of genomic data with other privacy-sensitive data (e.g. location, ancestry and other online social network – OSN – data) increases privacy risk through the potential for cross-layer attacks.

Techniques from Information security to protect genomic data

Information security is the practice of protecting information from unauthorized access, use, disclosure, disruption, modification, inspection, recording, or destruction [12]. Traditionally, information security has been widely used by governmental, military, and financial institutions. Only in the last few years has the field of information security grown and evolved significantly also in the medical context.
Tools protecting informational privacy by eliminating or minimizing personal data without the loss of the functionality of the information system are usually called privacy-enhancing technologies (PET) [13]. PET generally protect users' privacy by either breaking the link between individuals' identities and the data they provide (e.g., removing user's identities from published data), or by decreasing the amount of provided information (e.g., by using cryptographic tools or obfuscation techniques).
The idea of using technical solutions to guarantee the privacy of genomic data raises interesting debates. On one hand, the potential of genomic data for mankind is tremendous. Privacy-enhancing technologies can be considered as an obstacle to achieve these goals. On the other hand, to expedite advances in personalized medicine, genome-phenome association studies often require the participation of a large number of research participants. To encourage individuals to enroll in such studies, it is crucial to adhere to ethical principles, such as autonomy, reciprocity and, more generally, trust (e.g., guarantee that genomic data will not be misused). Technological solutions for genome privacy can be achieved by various techniques, such as

access control, cryptography, anonymization, or obfuscation.

Access control is the selective restriction of access to sensitive information by authorized people, whereas cryptography transforms usable information into a form that renders it unusable by anyone not having the decryption key. In modern cryptography, the distinction is made between symmetric encryption (e.g., for the encryption of files) (Fig. 1a), and (ii) asymmetric encryption (e.g., for digital signatures, for example to guarantee the authenticity of a piece of code) (Fig. 1b). Homomorphic encryption (Fig. 1c) is a special case of asymmetric cryptography.

Anonymization and obfuscation are two different privacy-preserving techniques. On one hand, anonymization is the process of removing tracks, or the electronic trail, on the data that would lead an eavesdropper to its origins, whereas on the other hand obfuscation deliberately muddles the data in order to prevent an attacker from interpreting it. Some examples of anonymization are k-anonymity and l-diversity (Fig. 2).

Cryptographic techniques typically reduce the efficiency of the algorithms, introducing storage and computational overhead, while preventing the users from «viewing» the data. Obfuscation-based methods reduce the accuracy (or utility) of genomic data. Therefore, especially when human life is at stake, using such privacy-enhancing techniques for genomic data can be delicate.

We need techniques that guarantee the security and privacy of genomic data, without significantly degrading the efficiency of the use of genomic data in research and healthcare. Developing PETs for genomic data presents challenges due to the architecture of the human ge-

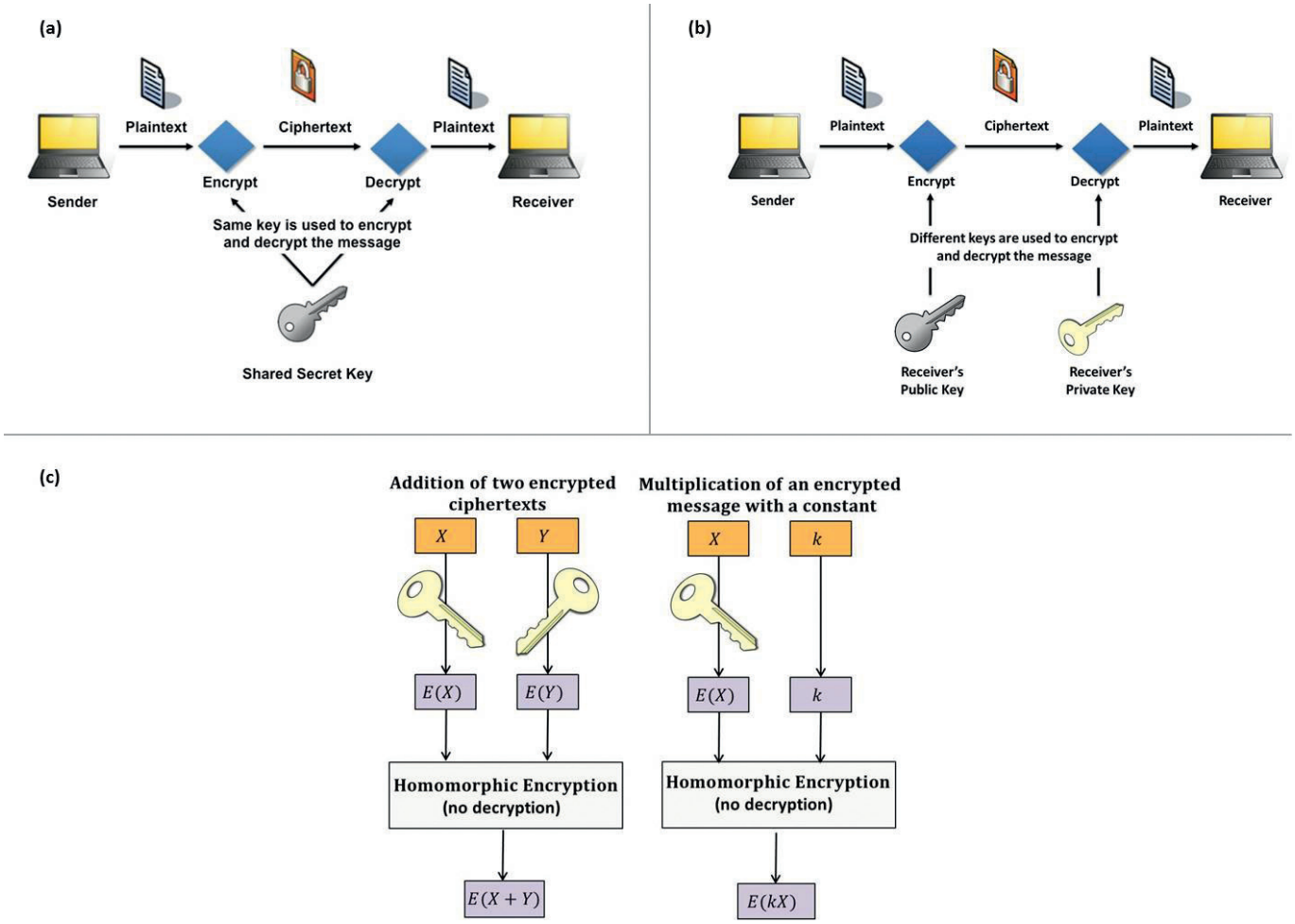

Fig. 1: a) Symmetric encryption: The same key is shared between the *Sender* and the *Receiver* and it is used to encrypt and decrypt the message. **b)** Asymmetric encryption: Different keys are used to encrypt and decrypt the message. The *Receiver (R)* is provided with a pair of keys composed by a Public Key and a Private Key. The *Sender* uses R's Public Key to encrypt the plaintext and R uses his Private key to decrypt. **c)** Homomorphic encryption is a form of encryption that allows specific computations on the ciphertext domain. For example, **(i)** the product of two ciphertexts is equal to the encryption of the sum of their corresponding plaintexts, and **(ii)** a ciphertext raised to a constant number is equal to the encryption of the product of the corresponding plaintext and the constant.

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	1004	27	Swiss	Heart Disease
2	1004	25	German	Heart Disease
3	1003	28	France	HIV
4	1003	26	Italian	Cancer
5	1100	54	Indian	Cancer
6	1100	51	Russian	Heart Disease
7	1111	48	Swiss	HIV
8	1111	47	Swiss	HIV
9	1010	34	Swiss	Alzheimer's
10	1010	39	Italian	Alzheimer's
11	1011	31	Greek	Alzheimer's
12	1011	36	American	Alzheimer's

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	10**	< 30	*	Heart Disease
2	10**	< 30	*	Heart Disease
3	10**	< 30	*	HIV
4	10**	< 30	*	Cancer
5	10**	3*	*	Alzheimer's
6	10**	3*	*	Alzheimer's
7	10**	3*	*	Alzheimer's
8	10**	3*	*	Alzheimer's
9	11**	≥ 40	*	Cancer
10	11**	≥ 40	*	Heart Disease
11	11**	≥ 40	*	HIV
12	11**	≥ 40	*	HIV

	Non-Sensitive			Sensitive
	Zip Code	Age	Nationality	Condition
1	10**	< 35	*	Heart Disease
2	10**	< 35	*	Heart Disease
3	10**	< 35	*	HIV
4	10**	< 35	*	Cancer
5	10**	< 35	*	Alzheimer's
6	10**	< 35	*	Alzheimer's
7	10**	≥ 35	*	Cancer
8	10**	≥ 35	*	Heart Disease
9	10**	≥ 35	*	HIV
10	10**	≥ 35	*	HIV
11	10**	≥ 35	*	Alzheimer's
12	10**	≥ 35	*	Alzheimer's

(a) (b)

Fig. 2: Example of anonymization: In **(a)**, the original table has been modified to fulfill *k-anonymity*: each released record is now indistinguishable from at least (k-1) others on its quasi-identifier attributes (in this case the zip code, age, and nationality). The table is further anonymized in **(b)**, to fulfill *l-diversity*: each group of tuples sharing the same quasi-identifier must have at least l distinct sensitive values which are roughly of equal proportions. For example, if an attacker knows that a given patient, aged 46, is in the database, neither from table **(a)** nor from table **(b)** can he figure out the clinical condition of the patient. However, if that patient is 36, from **(a)** the attacker can guess the patient has Alzheimer's, whereas no conclusion can be drawn from **(b)**.

nome and to the fast evolving knowledge in the field of genomics.

Existing works for genome protection

A first example (Fig. 3) of genome protection is the protection of genomic privacy in medical tests and personalized medicine. In particular, in our research effort [14,15], we focus on genetic disease susceptibility tests by developing a new architecture between the patient and the medical unit. We make use of homomorphic encryption and proxy re-encryption, and propose a privacy-preserving disease susceptibility test. Assuming the whole genome sequencing (WGS) to be done by a certified institution, we propose to store patients' genomic data encrypted at a «storage and processing unit» (SPU). Our architecture, while preserving the privacy of patients' genomic data, enables the medical unit to retrieve the encrypted genomic data from the SPU and process it for medical tests and personalized medicine methods.

The second system (Fig. 4) that we propose [16] is meant to protect the privacy of aligned, raw genomic data. Geneticists usually prefer to store this particular format of genomic data in addition to individuals' variation profiles (compact and summarized form of the raw data already protected in the previous framework), mainly because of the immaturity of bioinformatics algorithms and sequencing platforms. The raw genomic data of a patient includes mil-

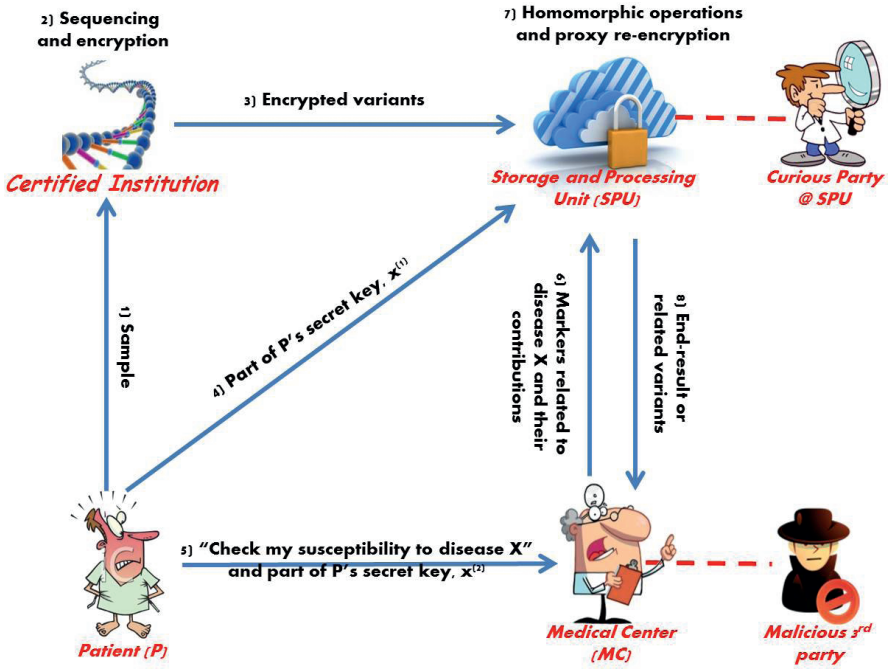

Fig. 3: Proposed solution: **(0)** The cryptographic keys (public and secret keys) are generated and distributed to the patients; **(1)** Patient (P) provides his sample (e.g., his saliva) to a certified institution (CI) for DNA extraction and sequencing; **(2)** After the sequencing, the CI encrypts P's genetic variants with its public key; **(3)** CI send the encrypted variants to the storage and processing unit (SPU); **(4)** P sends part of his secret key to SPU; **(5)** P asks to the medical center (MC) to check his susceptibility for a certain disease *X* and sends the remaining part of his secret key; **(6)** MC sends to SPU the markers identifiers related to disease *X* and their contribution (a genetic markers can influence the risk for a disease in different ways); **(7)** SPU, through homomorphic operations, computes the encrypted susceptibility and partially decrypts the result with its part of the secret key; **(8)** The partially decrypted end-result is sent back to the MC which performs the final decryption and gets the plaintext end-result. Even if the MC is a malicious party, it can only see the final result of the computation and can infer nothing about P's genome.

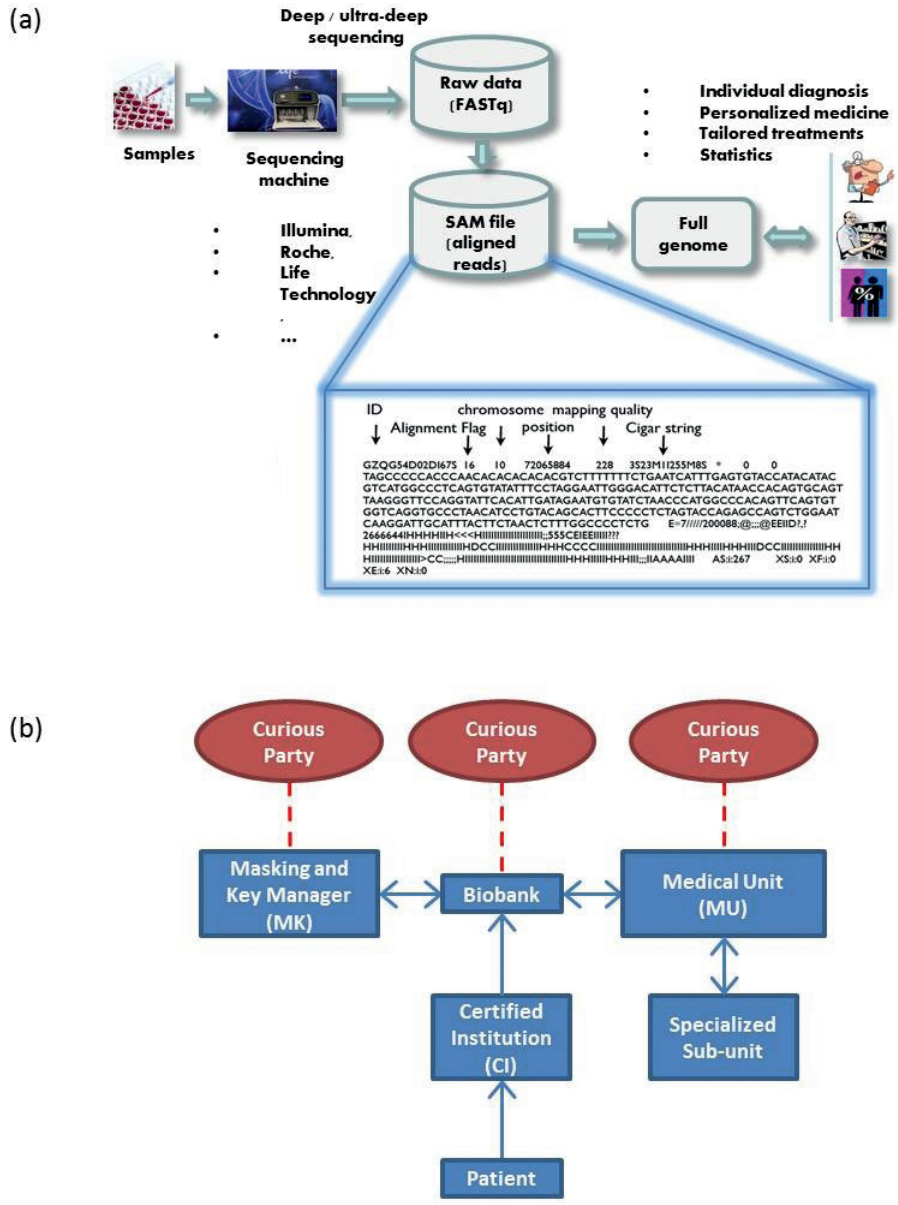

Fig. 4: a) Flow of the genomic data from the biological sample to the fully digitized genome. **b)** Schema of the proposed framework: SAM files are stored in an encrypted fashion at the *Biobank* by using pseudonyms; *Medical Units (MU)* need to be approved institutions in order to request ranges of nucleotides; the cryptographic keys of the patients are stored at the *Masking and Key Manager (MK)*.

lions of short reads of 100 to 400 nucleotides. Our proposed scheme stores these short reads at a biobank in encrypted form and enables a medical unit to privately retrieve a subset of the short reads of the patients without revealing the nature of the genetic test to the biobank. Furthermore, this architecture enables the biobank to mask particular parts of the retrieved short reads if (i) some parts of the provided short reads are out of the requested range, or (ii) the patient does not give consent to some parts of the provided short reads (e.g., parts revealing sensitive diseases) (Fig. 5).

Several other research groups have tackled various aspects of genome privacy, including (i) private string searching and comparison on the DNA sequence [17,18], (ii) private release of aggregate data [19–21], and (iii) private clinical genomics [22]. More generally, the protection of medical data has been very actively addressed by researchers. Many ad-hoc electronic health record (EHR) systems use cryptographic protocols to store medical information in a secure fashion and to define the access rights of the medical units [23,24]. Yet, research in the genomic privacy field is still in its infancy.

Kin genomic privacy

Another important aspect of genomic privacy is kin privacy. Even if one person's genome is not disclosed, a family member's genome can leak significant information about that person. Some believe that they have nothing to hide about their genetic structure; hence they might decide to give full consent for the publication of their genomes on the Internet to help genomic research. However, our DNA sequences are highly correlated to our relatives' sequences. Consequently, a person revealing his genome not only damages his own genomic privacy, but also puts his relatives' privacy at risk [25]. Currently, people do not need consent from their relatives to share their genome online.

Recently, our research group has provided a method to estimate the genome privacy loss of an individual when the genomes of some of his or her relatives are publicly available [26].

Conclusion

Genome sequencing is an emerging phenomenon and, as it is radically novel, it brings both great medical opportunities and significant privacy concerns. Genomic data will be increasingly available on the Internet, especially in citizen-contributed environments, (e.g., online social networks). Sharing this extremely sensitive data will raise unprecedented privacy concerns, because the extent of the information that can be revealed by a genome and how it could be used is yet unknown.

Key messages

- Impressive advances in genome sequencing have paved the way to a variety of revolutionary applications in modern healthcare.
- However, because of the genome's highly sensitive nature, this progress raises important privacy and ethical concerns that need to be addressed.
- Therefore, there is a clear need to support personalized medicine, genome research, forensic investigation, and direct-to-consumer genomics, while respecting privacy. Information security techniques will play an instrumental role in this process.
- Yet, it is important to observe that technical solutions are often not enough for protecting genomic privacy, therefore legal and professional guidelines are needed to govern how genomic information is handled by the different stakeholders.

Key messages

- Eindrückliche Fortschritte in der Genom-Sequenzierung haben den Weg für viele bahnbrechende Anwendungen in der modernen Medizin geebnet.
- Weil das Genom sehr sensitive Informationen über dessen Besitzer enthält, müssen auch Bedenken bezüglich Ethik und Datenschutz Rechnung getragen werden.
- Personalisierte Medizin, Genom-Forschung, forensische Forschung sowie die Direct-to-Consumer-Genomik sind zu unterstützen, Datenschutz und Privatsphäre müssen jedoch gewahrt sein. Techniken für Informationssischerheit werden in diesem Feld eine grössere Rolle spielen.
- Die technischen Lösungen sind jedoch häufig nicht ausreichend, um die «Genom-Privacy» zu gewährleisten. Deshalb braucht es für alle Beteilgten rechtliche und fachliche Guidelines für den Umgang mit genomischen Informationen.

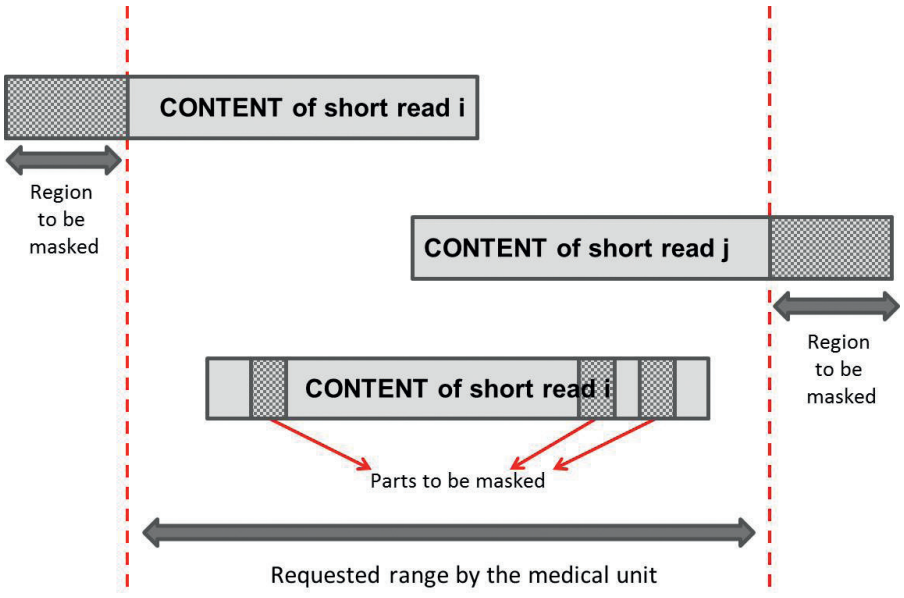

Fig. 5: The masking technique: Preventing the leakage of extra information in the short read (SR) to the MU by masking the positions of the SR external to the requested range and «non-consented» nucleotides by the patient. The encrypted SR are masked at the biobank.

Therefore, there is a clear need to support personalized medicine, genome research, forensic investigation, and direct-to-consumer genomics, while respecting privacy. Information security techniques will play an instrumental role in this process. An extensive overview of our research endeavors is accessible through our Web site [27] and, recently, a community platform [28] has also been developed for sharing information about the protection of genomic privacy. It is important, however, to observe that technical solutions are often not sufficient for protecting genomic privacy, therefore legal and professional guidelines are needed to govern how genomic information is transmitted, stored and processed by the different stakeholders. Mitigating privacy issues will require long-term collaboration among geneticists, other healthcare providers, ethicists, lawmakers, and computer scientists.

Zusammenfassung

Vielen Wissenschaftlern und Medizinern zufolge wird die Genomik auch in Zukunft eine grosse Rolle in der Medizin spielen. Eindrückliche Fortschritte im Bereich der Genom-Sequenzierung haben den Weg für bahnbrechende Anwendungen in der modernen Medizin geebnet. Besonders das wachsende Verständnis über das menschliche Genom und dessen Zusammenspiel mit Krankheiten und Therapieansprechen versprechen Verbesserungen für die Prävention und die personalisierte Medizin. Damit treten jedoch auch ethische Bedenken und Fragen bezüglich Privatsphäre bzw. Datenschutz auf den Plan. Denn jedes Genom ist ein Identifikator seines Besitzers und enthält daher sehr persönliche und Privatsphäre-sensitive Daten. Dieser Artikel beleuchtet wichtige Themen und Methoden bezüglich Privatsphäre

und Datenschutz im Zusammenhang mit Genomdaten.
Schlüsselwörter: Genom – Privatsphäre – Datensicherheit – Kryptographie Anonymisierung

Résumé

Selon l'avis de beaucoup de scientifiques et de cliniciens la génomique est appelé à jouer un rôle clé dans le domaine de la médecine. Des progrès impressionnants dans le séquençage du génome ont ouvert la voie à une variété d'applications révolutionnaires dans les soins de santé. En particulier, la compréhension croissante du génome humain, de sa relation aux maladies et la réponse aux traitements promet d'améliorer la médecine préventive et personnalisée. Néanmoins, ce progrès est source d'une inquiétude sur le plan éthique et de la protection des données, une inquiétude qui doit être discutée. En effet, chaque génome représente le moyen ultime d'identifier la personne dont il provient et, contient dès lors des données personnelles hautement sensibles sur le plan privé. Dans cet article, après avoir résumé les progrès récents en génomique, nous allons discuter de quelques points importants concernant la politique de confidentialité sur le plan de l'information génomique humaine, de même que des méthodes mises en place pour les gérer.
Mots-clés: génome – confidentialité des données – sécurité de l'information – anonymisation – cryptographie

Correspondence address

Prof. Dr. Jean-Pierre Hubaux
EPFL IC ISC LCA1
BC 207 (Bâtiment BC)
Station 14
1015 Lausanne

Jean-Pierre.Hubaux@epfl.ch

Bibliography

1. Ayday E, De Cristofaro E, Hubaux JP, Tsudik G: The Chills and Thrills of Whole Genome Sequencing. IEEE Computer Magazine 2014.
2. Malin BA: An evaluation of the current state of genomic data privacy protection technology and a roadmap for the future. J Am Med Inform Assoc 2005; 12: 28–34.
3. Homer N, Szelinger S, Redman M, et al.: Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS Genetics 2008; 4: e1000167.
4. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y: Identifying personal genomes by surname inference. Science 2013; 339: 321–324.
5. Venter J, Adams MD, Myers EW, et al.: The sequence of the human genome. Science 2001; 291: 1304–1351.
6. Botstein D, Risch N: Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nature Genetics 2003; 33: 228–237.
7. MonDossierMedical.ch. http://mondossiermedical.ch/; visited on the 25/03/2014.
8. healthbank. http://healthbank.ch/; visited on the 25/03/2014.
9. Brunham L, Hayden M: Whole-genome sequencing: the new standard of care? Science 2012; 336: 1112–1113.
10. 23andMe. https://www.23andme.com/; visited on the 25/03/2014.
11. Counsyl. https://www.counsyl.com/; visited on the 25/03/2014.
12. Cherdantseva Y, Hilton J: Information Security and Information Assurance: The Discussion about the Meaning, Scope and Goals. Organizational, Legal, and Technological Dimensions of Information System Administrator. IGI Global Publishing 2013; 167–198.
13. van Blarkom GW, Borking JJ, Olk JGE: Handbook of Privacy and Privacy-Enhancing Technologies. (The Case of Intelligent Software Agents). Private Incorporated Software Agent (PISA): 2003.
14. Ayday E, Raisaro JL, McLaren PJ, Fellay J, Hubaux JP: Privacy-Preserving Computation of Disease Risk by Using Genomic, Clinical, and Environmental Data. USENIX Security Workshop on Health Information Technologies (HealthTech ,13), Washington, D.C. 2013.
15. Ayday E, Raisaro JL, Rougemont J, Hubaux JP: Protecting and Evaluating Genomic Privacy in Medical Tests and Personalized Medicine. ACM Workshop on Privacy in the Electronic Society (WPES 2013), Nov. 2013, Berlin, Germany.
16. Ayday E, Raisaro JL, Hengartner U, Molyneaux A, Hubaux JP: Privacy-Preserving Processing of Raw Genomic Data. 8th Data Privacy Management (DPM 2013) International Workshop (in conjunction with ESORICS 2013), Sep. 2013, Egham, UK.
17. Troncoso-Pastoriza JR, Katzenbeisser S, Celik M: Privacy preserving error resilient DNA searching through oblivious automata. In Proceedings of the 14th ACM Conference on Computer and Communications Security (CCS'07) Oct. 2007, Alexandria, VA, USA: 519–528.
18. Blanton M, Aliasgari M: Secure outsourcing of DNA searching via finite automata. In Proceedings of the 24th Annual IFIP WG 11.3 Working Conference on Data and Applications Security and Privacy (DBSec'10), Jun. 2010, Rome, Italy; 49–64.
19. Kantarcioglu M, Jiang W, Liu Y, Malin B: A cryptographic approach to securely share and query genomic sequences. IEEE Transactions on Information Technology in Biomedicine 2008; 12: 606–617.
20. Fienberg SE, Slavkovic A, Uhler C: Privacy preserving GWAS data sharing. In Proceedings of the IEEE 11th International Conference on Data Mining Workshops (ICDMW'11), Dec. 2011, Vancouver, Canada; 628–635.
21. Johnson A, Shmatikov V: Privacy-preserving data exploration in genome-wide association studies. In Proceedings of the 19th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD'13) Aug. 2013, Chicago, IL, USA; 1079–1087.
22. Baldi P, Baronio R, De Cristofaro E, Gasti P, Tsudik G: Countering gattaca: efficient and secure testing of fully-sequenced human genomes. In Proceedings of the 18th ACM conference on Computer and communications security, (CCS '11) Oct. 2011, Chicago, IL, USA; 691–702.

23. Narayan S, Gagne M, Safavi-Naini R: Privacy preserving EHR system using attribute-based infrastructure. In Proceedings of the 2010 ACM Workshop on Cloud Computing Security (CCSW'10), Oct. 2010, Chicago, IL, USA; 47–52.
24. Alshehri S, Radziszowski S, Raj R: Secure access for healthcare data in the cloud using ciphertext-policy attribute-based encryption. In Proceedings of the 28th International Conference on Data Engineering Workshops (ICDEW'12), Washington, DC, USA, 2012: 143–14.
25. Stajano F, Bianchi L, Liò P, Kor D: Forensic genomics: kin privacy, driftnets and other open questions. In Proceedings of the 7th ACM workshop on Privacy in the Electronic Society (WPES'08), New York, NY, USA, Oct. 2008; 15–22.
26. Humbert M, Ayday E, Hubaux JP, Telenti A: Addressing the concerns of the lacks family: Quantification of kin genomic privacy. In Proceedings of the 20th ACM Conference on Computer and Communications Security (CCS'13) Nov. 2013, Berlin, Germany; 1141–1152.
27. EFPL – LCA1 Genomic Privacy: http://lca.epfl.ch/projects/genomic-privacy/; visited on the 25/03/2014.
28. GenomePrivacy.org: https://genomeprivacy.org; visited on the 25/03/2014.