Poster: Membership Inference Against DNA Methylation Databases: Attacks and Defenses

Inken Hagestedt*, Mathias Humbert†, Pascal Berrang*, Irina Lehmann‡, Roland Eils§, Michael Backes*, Yang Zhang*
*CISPA Helmholtz Center for Information Security, firstname.lastname@cispa.saarland, backes@cispa.saarland
†Swiss Data Science Center, ETH Zurich and EPFL, mathias.humbert@epfl.ch
‡ Helmholtz Centre for Environmental Research Leipzig, UFZ, Leipzig, irina.lehmann@ufz.de
§ German Cancer Research Center (DKFZ), University of Heidelberg, r.eils@dkfz-heidelberg.de

***Abstract*—Access to biomedical data is key for the advancement of biomedical research. However, biomedical data raises new privacy concerns: membership inference attacks against a biomedical database can leak sensitive information, such as the participants' health status. In this paper, we study membership inference attacks on DNA methylation data, one of the most important epigenetic elements regulating the human health. We design three different types of attacks exploiting published summary statistics. Two of them are based on statistical tests and the third one on a machine learning model. Moreover, we exploit the dependencies between genome and methylation data to replace knowledge of the target's methylome by knowledge of the target's genome, which is currently more accessible. Our extensive evaluation shows that such membership inference attacks are effective. In order to mitigate these privacy risks, we rely on differential privacy and show that such defense is able to provide strong privacy guarantees at the cost of a significantly decreased utility. By restraining the number of released methylation regions to a few hundreds, we can reach an acceptable level of privacy without sacrificing all statistical utility.**

With the rapidly decreasing costs of molecular profiling, the types of available biomedical data are increasingly diverse and go beyond the genomes of individuals. DNA methylation is one of the most important new types of biomedical data. Being a key regulator of gene transcription, abnormal methylation patterns can lead to severe diseases, such as cancer [2]. Moreover, DNA methylation is also related to environmental cues, such as pollution, exposure to stress or cigarette smoke [8], [9], [1]. Despite being linked with such sensitive information, DNA methylation data is already available on various open research platforms, such as the Gene Expression Omnibus (GEO) [4]. Contrary to genomic data whose privacy has been extensively studied by the security research community [3], [7], [6], the privacy risks stemming from these more recent epigenomic data attracted less attention.

One of the most critical attacks in the biomedical research setting is membership inference, popularized by Homer et al.[5]: Given some raw data about a targeted individual, the attacker wants to know whether this individual is member of a dataset (i.e., has contributed his data) by solely relying on aggregated statistics about this dataset. Such a membership inference attack can have serious privacy implications if this dataset contains individuals carrying a specific disease.

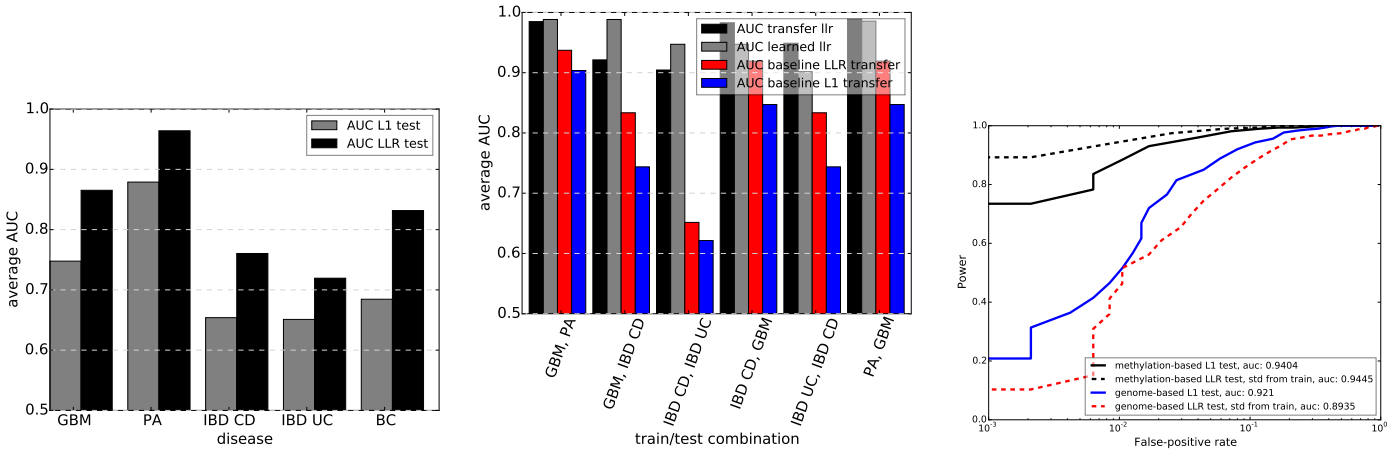
We aim at evaluating whether DNA methylation databases are also vulnerable to membership inference attacks. DNA methylation data is not only very sensitive as it can unveil severe diseases such as cancer, some regions of our methylation profiles are highly correlated with the genome, thus leakage of such data can indirectly expose family members' private data. As a consequence, anticipating privacy risks and mitigating them with technical means is of utmost importance.

a) Contributions: Specifically, we present multiple attacks against the membership privacy of individuals participating in DNA methylation-based studies. We consider two types of adversarial settings, both relying on mean DNA methylation statistics released about the databases. The first setting assumes the adversary to know its victim's DNA methylation profile, while the second setting assumes the victim's genome to be known instead.

For both adversarial settings, we design three types of membership inference attacks: one based on the L_1 distance, one based on the likelihood-ratio (LR) test and one based on a machine learning classifier trained on distance features. For the genome-based inference, we particularly design our attack to capture the probabilistic dependencies between the two types of biomedical data. We prove that the mean of the conditional distribution of the methylation values given the genomic values is a sufficient statistic for the genome-based attack.

We then conduct an extensive evaluation of our attacks on six diverse datasets, containing a total of 1,320 patients. Our results consistently demonstrate the success of this type of attack over different tissues and diseases. While the statistical test based on the LR test exceeds 0.7 AUC and reaches over 0.95 AUC in one case (see Figure 1a), machine-learning increases the AUC to over 0.9 in most cases. Additionally, the attacker's training data can be distinct from the target data, as our experiments on transferability demonstrate in Figure 1b. Even if the attacker only knows the target's genome, inference of the methylation values followed by a membership inference attack is possible, as Figure 1c shows.

Propelled by these results, we propose a differentially private mechanism. We empirically evaluate its effectiveness on our various datasets. While our mechanism is able to provide strong privacy guarantees, it also negatively affects the utility of the data. If the adversary gets access to the full set of methylation points, we cannot obtain perfect privacy and, at the same time, accurate statistics. However, if only a few hundred methylation values are released, the average amount of noise added decreases and reasonable privacy levels are reached, see Figure 2.

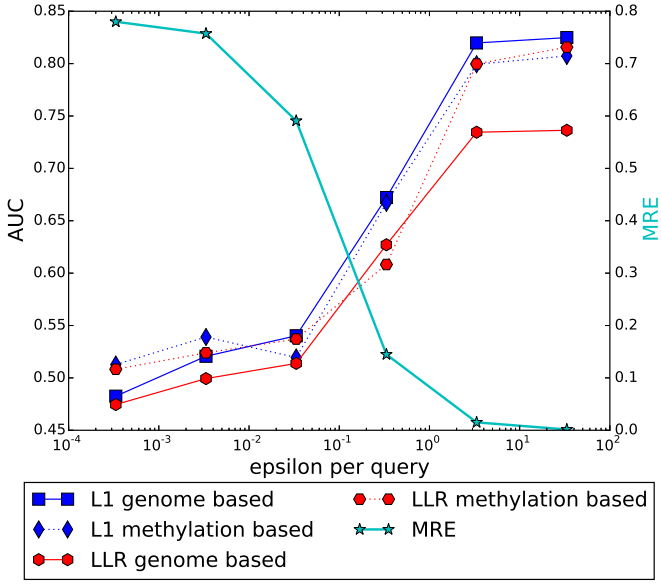


(a) AUC of the L_1 and LLR tests applied to all methylation positions.

(b) Transferability of learned models based on LLR features.

(c) ROC curves of methylation-based L_1 and LLR tests and genome-based L_1 and LLR tests.

Fig. 1: An overview on our attack evaluation results.



(a)

Fig. 2: Defense with Laplace noise and release of only a few hundred methylation values.

References

[1] T. Bauer, S. Trump, N. Ishaque, L. Thu rmann, L. Gu, M. Bauer, M. Bieg, Z. Gu, D. Weichenhan, J.-P. Mallm, S. Ro der, G. Herberth, E. Takada, O. Mu cke, M. Winter, K. M. Junge, K. Gru tzmann, U. Rolle-Kampczyk, Q. Wang, C. Lawerenz, M. Borte, T. Polte, M. Schlesner, M. Schanne, S. Wiemann, C. Geo rg, H. G. Stunnenberg, C. Plass, K. Rippe, J. Mizuguchi, C. Herrmann, R. Eils, and I. Lehmann, "Environment-induced epigenetic reprogramming in genomic regulatory elements in smoking mothers and their children," *Molecular Systems Biology*, vol. 12, no. 3, pp. 861–861, mar 2016.

[2] P. M. Das and R. Singal, "DNA methylation and cancer," *Journal of clinical oncology*, vol. 22, no. 22, pp. 4632–4642, 2004.

[3] Y. Erlich and A. Narayanan, "Routes for breaching and protecting genetic privacy," *Nature reviews. Genetics*, vol. 15, no. 6, p. 409, 2014.

[4] "Gene expression omnibus," https://www.ncbi.nlm.nih.gov/geo, accessed: 2017-20-07.

[5] N. Homer, S. Szelinger, M. Redman, D. Duggan, W. Tembe, J. Muehling, J. V. Pearson, D. A. Stephan, S. F. Nelson, and D. W. Craig, "Resolving individuals contributing trace amounts of dna to highly complex mixtures using high-density snp genotyping microarrays," *PLoS Genet*, vol. 4, no. 8, p. e1000167, 2008.

[6] A. Mittos, B. Malin, and E. De Cristofaro, "Systematizing genomic privacy research–a critical analysis," *arXiv preprint arXiv:1712.02193*, 2017.

[7] 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 (CSUR), 2015*, 2015.

[8] S. Trump, M. Bieg, Z. Gu, L. Thürmann, T. Bauer, M. Bauer, N. Ishaque, S. Röder, L. Gu, G. Herberth, C. Lawerenz, M. Borte, M. Schlesner, C. Plass, N. Diessl, M. Eszlinger, O. Mücke, H.-D. Elvers, D. K. Wissenbach, M. von Bergen, C. Herrmann, D. Weichenhan, R. J. Wright, I. Lehmann, and R. Eils, "Prenatal maternal stress and wheeze in children: novel insights into epigenetic regulation," *Scientific Reports*, vol. 6, p. 28616, jun 2016.

[9] L. G. Tsaprouni, T.-P. Yang, J. Bell, K. J. Dick, S. Kanoni, J. Nisbet, A. Viñuela, E. Grundberg, C. P. Nelson, E. Meduri, A. Buil, F. Cambien, C. Hengstenberg, J. Erdmann, H. Schunkert, A. H. Goodall, W. H. Ouwehand, E. Dermitzakis, T. D. Spector, N. J. Samani, and P. Deloukas, "Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation," *Epigenetics*, no. December, pp. 00–00, oct 2014.

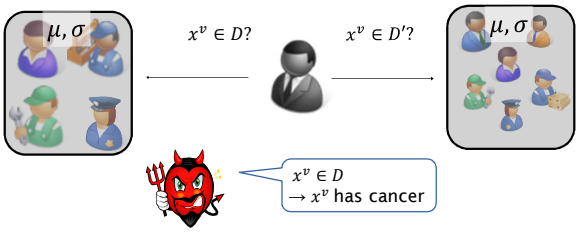
Membership Inference Against DNA Methylation Databases: Attacks and Defenses

Inken Hagestedt[1], Mathias Humbert[2], Pascal Berrang[1], Irina Lehmann[3], Roland Elis[4], Michael Backes[1], Yang Zhang[1]

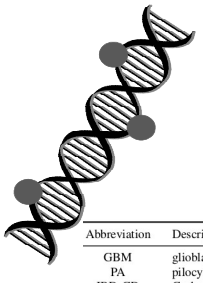
[1]CISPA Helmholtz Center for Information Security [2]Swiss Data Science Center [3]Helmholtz Centre for Environmental Research Leipzig, UFZ, [4]German Cancer Research Center, DKFZ

The Question

Are membership inference attacks possible given only mean μ and standard deviation σ?



The Data



- DNA methylation: additional molecule (methyl group) attached to DNA
- represented as value in [0, 1]
- methylation patterns vary between tissues, due to environmental factors and due to diseases

Abbreviation	Description	Tissue Type	Number of Patients	GSE identifier
GBM	glioblastoma	brain cancer	136	GSE36278
PA	pilocytic astrocytoma	brain cancer	61	GSE44684
IBD CD	Crohn's disease	blood	77	GSE87640
IBD UC	ulcerative colitits	blood	79	GSE87640
BC	breast cancer	breast cancer	892	-[3]
WGBS	genome and methylation data	blood	75	not publicly available

Statistics-based Attack

→ two statistical tests:

mean only:

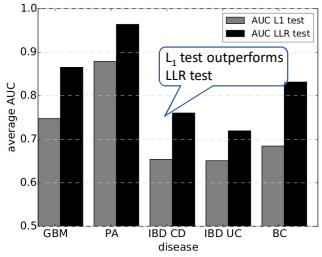
$$L_1(x^j) = |x^j - \mu_r^j| - |x^j - \mu_c^j|$$

$$L_1(x) = \underset{j \in 0,\dots,n}{\text{ttest}}\ L_1(x^j)$$

mean reference; mean case; combination of all methylation values with students t-test

mean and standard deviation:

$$LLR(x) = \sum_{i=0}^{n} \frac{(x^j - \mu_r^j)^2 - (x^j - \mu_c^j)^2}{2\,(\sigma^j)^2}$$



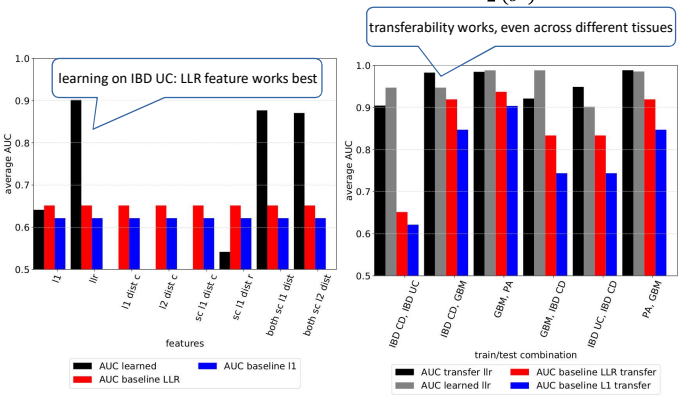
ML-based Attack

→ learn which distance magnitude is informative

distance features: $|x^j - \mu_c^j|$ $(x^j - \mu_c^j)^2$

scaled versions: $\frac{|x^j - \mu_c^j|}{\sigma^j}$ $\frac{(x^j - \mu_c^j)^2}{(\sigma^j)^2}$

test-inspired features: $|x^j - \mu_c^j| - |x^j - \mu_r^j|$ $\frac{(x^j - \mu_c^j)^2 - (x^j - \mu_r^j)^2}{2\,(\sigma^j)^2}$



Genome-based Attack

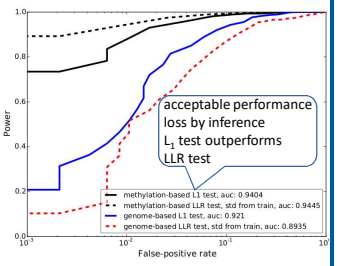
→ exploit correlation between genome and methylation:

$$f_g(x^j) = p(X_j = x^j \mid G = g) = \frac{1}{\sqrt{2\pi\sigma_{j,g}}} \exp\left(\frac{-(x^j - \mu_{j,g})^2}{(\sigma_{j,g})^2}\right)$$

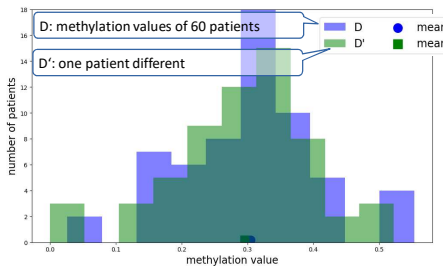
probability of methylation value modeled by Gaussian

same LLR test on all related positions: $LLR(g) = \sum_{j=1}^{m_c} \frac{(\mu_{j,g} - \mu_r^j)^2 - (\mu_{j,g} - \mu_c^j)^2}{2\,(\sigma^j)^2}$

expected methylation value given the genome



Defense with Differential Privacy



difference informative

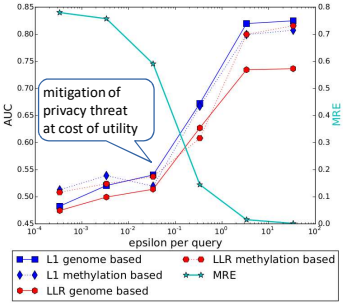
→ output a random mean

that hides the contribution of the changed entry

formally: $\Pr[M(mean(D)) = \mu] \leq e^{\epsilon} \Pr[M(mean(D')) = \mu]$

where $M(mean(D)) = mean(D) + Lap\left(\frac{m}{n\epsilon}\right)$

privacy parameter; number of positions here: 300.000; number of patients here: 60



CISPA HELMHOLTZ CENTER FOR INFORMATION SECURITY

SDSC

HELMHOLTZ ZENTRUM FÜR UMWELTFORSCHUNG UFZ

dkfz. GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION

@ inken.hagestedt@cispa.saarland
yang.zhang@cispa.saarland