Towards Privacy Compliance in the Management of Raw Genomic Data

Erman Ayday¹, Jean Louis Raisaro¹, Urs Hengartner², Zhicong Huang¹, Adam Molyneaux³, Jean-Pierre Hubaux¹ ¹{firstname.lastname@epfl.ch}, ²urs.hengartner@uwaterloo.ca, ³amolyneaux@sophiagenetics.com

- . . .



													E>	()	n	η		e										
Position	1	2	3	4	5	6	7	8	9	10	11	12	2 13	8 14	l 15	16	17	,		18	19	20	21	22	23	24	25	2
Reference	Α	G	С	Α	Т	G	Т	Т	Α	G	Α	T	· A	Α	G	Α	Т	*	*	Α	G	С	Т	G	Т	G	С	•
Content of the SR									a	а	a	(A	Α	*	A	T	G	С	•	•		Т	Α	т	G	С	
														,		\		· · · ·	,		\downarrow						/	
Ciga	r S	Stri	ng	(C	S)			⇒		35	5		31	M	10) 2	Μ	2	21		3N					8	M	

Erman Ayday, Jean Louis Raisaro, Urs Hengartner, Adam Molyneaux, and Jean-Pierre Hubaux. Privacy-Preserving Processing of Raw Genomic Data. 8th Data Privacy Management (DPM 2013) International Workshop (in conjunction with ESORICS 2013), Egham, UK, Sep. 2013. Web link: http://lca.epfl.ch/projects/genomic-privacy/



How the leakage of genomic data from short reads threatens the genomic privacy of a patient?

Without masking in place:



and SNPs is 0!!

Implementation and Complexity



Evaluation

Genomic data leakage of the patient as a result of various disease susceptibility tests each requiring a different number of SNPs from different parts of the patient's genome.

DISEASE TESTED	LEAKED SNP	NATURE OF THE LEAKED SNP							
	'rs1799724'	Susceptibility to Vascular Dementia							
	'rs6265'	Susceptibility to Memory Impairment							
	'rs6265'	Body Mass Index							
	'rs6265'	Smoking behavior							
	'rs6265'	Weight							
	'rs669'	Alpha-2-Macroglobulin Polymorphism							
Alzheimer's Disease	'rs429358'	Stroke							
	'rs429358'	Hyperlipoproteinemia type 3							
	'rs429358'	Brain Imaging							
	'rs4420638'	Total Cholesterol							
	'rs4420638'	HDL Cholesterol							
	'rs4420638'	LDL Cholesterol							
	'rs4420638'	Longevity							
	'rs4420638'	Coronary Artery Disease							
	'rs4420638'	Type II Diabetes							