# A fruitful transfer model of genomic data to a common language

### Dr. Francesco Proietti CASPUR

3RD ANNUAL COURSE OF PHARMACOGENETICS AND PERSONALIZED MEDICINE Workshop: Genomics contribute in wellness, fitness and sport



15/01/12 Machine to read individual's DNA for \$1,000 - FT.com
FINANCIAL TIMES

January 10, 2012 5:06 am

### Machine to read individual's DNA for \$1,000

The next generation sequencing technology is quickly evolving

By Clive Cookson

A US biotechnology company will on Tuesday announce the first machine that can read all 3bn letters of an individual's DNA for as little as \$1,000 – a development that will greatly accelerate medical treatment tailored to a patient's genes but also raises ethical questions.

#### Soon, \$1,000 Will Map Your Genes



#### ByRON WINSLOW And SHIRLEY S. WANG

SAN FRANCISCO—The quest to harness the power of DNA to develop personalized medicine is on the threshold of a major milestone: the \$1,000 genome sequencing.

Life Technologies Corp., a Carlsbad, Calif., genomics company, plans to introduce Tuesday a machine it says will be able to map an individual's entire genetic makeup for \$1,000 by the end of this year. Moreover, the machine and accompanying microchip technology, both developed by the company's lon Torrent unit, will deliver the information in a day, the company says.



If Life Technologies delivers on the claim, it would likely make the company the first among a group of rivals racing to produce a \$1,000 gene map. The current cheapest sequencing costs about \$3,000 and takes a week.

The goal, triggered in part by an initiative launched by the U.S. government's National Human Genome Research Institute in 2004, already has resulted in a

dramatic cost reduction in sequencing all three billion units of DNA, known as basepairs, that make up the human genetic code. The next generation sequencing technology is quickly evolving



CORRIERE DELLA SERA	Data 11-01-2012 Pagina 33 Foglio 1/2
Salute e tecnologia Presentato negli Stati Uniti lo strumento grande come una sta Mille dollari per farsi legge	impante. Dubbi dei bioetici Pre il Dna
Una macchina decifra il genoma in 24 ore. Spera	nze per nuove cure
MILANO — Leggere l'intero Intanto il genoma low cost o cerca un posto di lavoro.	corso della vita: il Dna non cam-

genoma di una persona costa, non sarà subito disponibile per oggi, mille dollari soltanto. Una il pubblico americano: la Food società americana di biotecno- and Drug Administration non logie, la Life Technologies, ha ha ancora preso posizione, ma appena presentato una macchi- in generale, per quanto riguarna in grado di decifrare, in me- da i test sul Dna, raccomanda no di 24 ore, i tre miliardi di let- cautela, perché una cattiva intere che compongono il Dna terpretazione dei dati può por-

taro a taranio coorrette e com-

della genziamento

hette a di-

mi di in-

The next generation sequencing technology is quickly evolving

É vero, poi, che conoscere la predisposizione a una determinata malattia (perché nel Dna è presente una certa variante. cioè una seguenza diversa rispetto al normale) può aiutare nella prevenzione, ma può esporre a esami inutili (con aumento dei costi per la sanità) o a un danno psicologico dell'individuo, se l'informazione è mal gestita.

Essere predisposti a una malattia (per esempio a un certo tibia, ma cambiano le interpretazioni che se ne danno.

É ipotizzabile, dunque, che il sequenziamento dell'intero genoma trovi oggi un'applicabilità reale soprattutto nel campo della ricerca. Per esempio, può essere utile per studiare persone affette da un tumore o dall'autismo o addirittura da malattie sconosciute, per cercare eventuali difetti genetici che ne stanno alla base. Può essere utilizzato (visto

The huge amount of data produced by the sequencing must be organized by means of bioinformatics analysis

> Only about 1% of the three billion letters in the human genome directly codes for proteins

Rome 9-10 February 2012



- Correlation of genetic, environmental, and lifestyle factors to actual illnesses can be difficult
  - Unknown or incompletely understood processes
  - Patient's body is not a closed, controlled laboratory
- Classic analytical and statistical approaches could be stressed by huge number of variables

Mithocondrial superoxide dismutase



Val16Ala polymorphism

Number of mutations	healthy people	people with asthenia	Tot
0 13 (26.5%)		36 (73.5%)	49
1	4 (7.7%)	48 (92.3 %)	52
2	5 (17.9%)	23 (82.1 %)	28
Tot	22 (17.0%)	107 (82.9%)	129

Even if this enzyme is not the only genetic factor, it probably plays a role in the onset of asthenia





- Self-Organizing Maps may help
  - Established in the 80s
  - A powerful technique to classify complex, unknown patterns
- SOM are modeled on efficient brain cortex functions that allow us to:
  - Recognize a schoolfellow after 30 years
  - Identify the original piece of music even if heavily rearranged
  - Discern casual coincidence from correlation

## Self-Organizing Maps in Action







#### Topology of mental disorders from symptoms



## Classification of astrophysics papers

Atomic D <mark>ata</mark>	Spe	ctro Abu	indance	s '	Var*		uvby	
•	•			•				
- Gai	l <mark>a</mark> xies	Models	A			UBV	-Ast	<mark>er</mark> oids
	• (•		. Lat	.e*	• (	LÜST	ERS	-
		QS0s	•		Open		. Ga	laxies
))))	•			•		•		•
	. Rad	<mark>io .</mark> X	- <mark>r</mark> ay	. В	inari	es.		-
	•			- G1	obula	r I	Astrom	etry
St <mark>ars</mark> ]	IR .	. Sp	Class	. C1	uster	s .		

## Matching CVs against job offerings

		Area jo	b placement		
	🚬 🐅 🔊 📰 🛄	ro	SAPIENZA	Contat	Powered by CASPUR
RICERCA LAVORO	RICERCA TIROCINI	RICERCA AZIENDE	AREA RISERVATA		LOG OUT
RICERCA OPPO	RTUNITÀ DI LAVO	DRO			
Cosa?		Dove?		CERC	A
Filtri: Retribuzion	ne 👻 Lingua	A	rea Professionale 👻	Mod. Lavorativa 👻	Rimuovi Filtri
≷isultati <b>1 - 10</b> di <b>82</b>				Ordina per	dice di prossimità 💌

12 JAN	Addetto/a Integrazioni e Progetti Speciali Area IT SEL 01/12 Cerved Group Spa ROMA // Capoprogetto informatico // DA 1000 A 1500 Euro // FULL-TIME // Scadenza: 27/01/2012 Descrizione Ruolo La risorsa sarà impegnata nello sviluppo informatico dei Progetti Speciali Cerved Group. Studierà la fattibilità pratica coniugando le informazioni ricevute dal proprio responsabile con le necessità dell'azienda nostra cliente. Si confronterà quindi telefonicamente, o all'occorrenza de visu, con il cliente. Il Candidato II candidato ideale ha 30-35 anni, laureato in Inge				
13 DEC	Indice di prossimità       70%         Sviluppatore       Gli studi descritti nel curriculum risultano non attinenti alle richieste dell'opportunità' di lavoro. Tuttavia il sistema neural ha rilevato una buona corrispondenza fra il testo del curriculum e quello dell'opportunità' di lavoro. Le esperienze lavorative e/o professioni desiderate elencate nel curriculum risultano attinenti all'area professionale cui si rivolge         MILANO // Analista programma       Gli studi descritti nel curriculum risultano non attinenti alle richieste dell'opportunità' di lavoro. Le esperienze lavorative e/o professioni desiderate elencate nel curriculum risultano attinenti all'area professionale cui si rivolge	/IE // Scadenza: 31/12/2012			
	Ai fini dell'ampliamento del suo organico, Kirey S.r.I. sta ricercando per la sua sede di Milano: • Laureandi/Neo laureati (laurea triennale/specialistica) in: Informatica, Ingegneria Informatica, Ingegneria Gestionale, Ingegneria Elettronica, Ingegneria delle Telecomunicazioni, Matematica e Fisica. Le risorse selezionate saranno inserite in un team di lavoro seguendo un percorso formativo e Int				
	Indice di prossimità 🔐 69%				

# Gene prediction in the *Brucella suis* genome



# Topology of mental disorders from symptoms



### SOMs for Patient Oriented Health Care

- SOM projects a space described by too many variables onto a twodimensional or one-dimensional space, preserving clusters of similar patterns
- Factors must be transformed or combined in proper descriptor i.e. variables in the original space
- Self-organization algorithm must then be applied until the map relaxes to equilibrium
- Clusters must be identified by testing against known associations of factors to illnesses
- A prudent addition of irrelevant or questioned factors may help
  - To better separate clusters
  - To make new discoveries
- The map can then be used to classify new patients' data