L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura

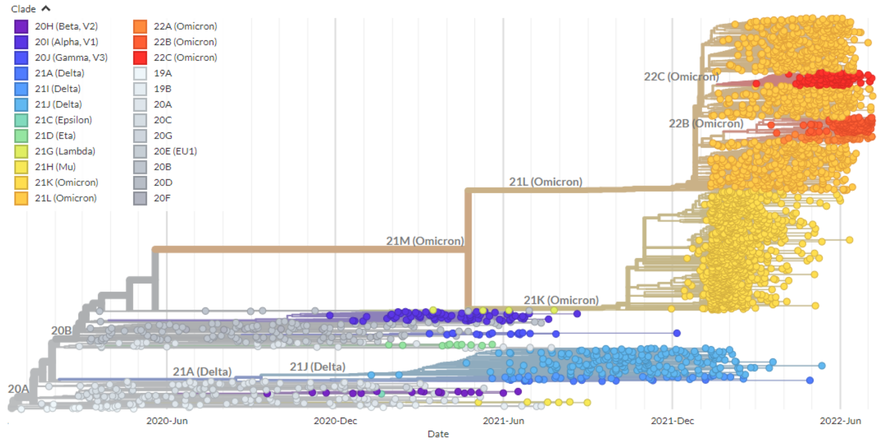
L'Intelligenza Artificiale per l'analisi dei genomi e delle mutazioni genetiche

Michele Morgante

IGA
ISTITUTO DI GENOMICA APPLICATA

L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

Analisi delle sequenze del genoma di SARS-Cov2: dagli approcci filogenetici al machine learning al deep learning

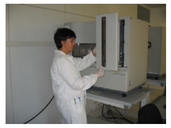


laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

- Prime applicazioni di machine learning in genetica molecolare: anni '80
- Identificazione di elementi nelle sequenze di DNA
  - Porzioni codificanti
  - Porzioni ripetute
- Accelerazione nell'uso di AI-metodi con l'avvento del Next Generation Sequencing
  - Enormi moli di dati prodotti
  - Tanti tipi diversi di dati prodotti con la stessa tecnologia

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

platform	ABI 3730	HiSeq2500	NovaSeq6000	NovaSeqXPlus
Method	Sanger	Illumina	Illumina	Illumina
Throughput/run	76.000 bp	1 Tbp	6 Tbp	16 Tbp
Throughput/day	1,824 Mbp	150.000 Mbp	3.000.000 Mbp	8.000.000 Mbp
Fold increase throughput	n.a.	82.000	1.640.000	4.386.000
Cost (Euro/Gbp)	1.250.000	30	12,5	2,5
Fold decrease cost	n.a.	42.000	100.000	500.000



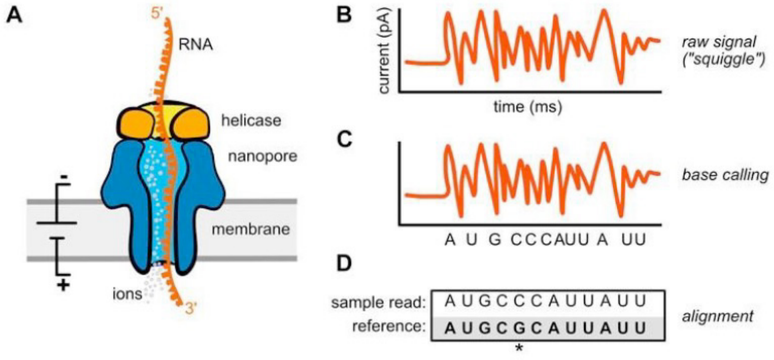

#sanita2030

THE NGS REVOLUTION

www.sanita2030.it

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

3rd generation sequencing: Nanopore sensing

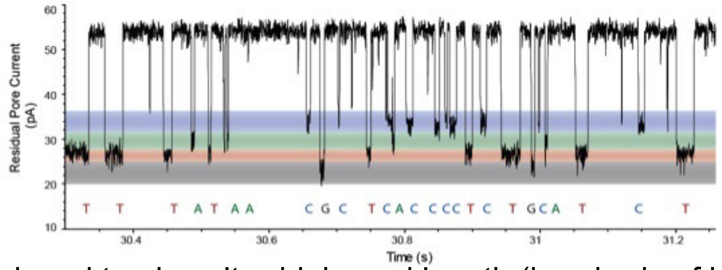


#sanita2030

www.sanita2030.it

L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

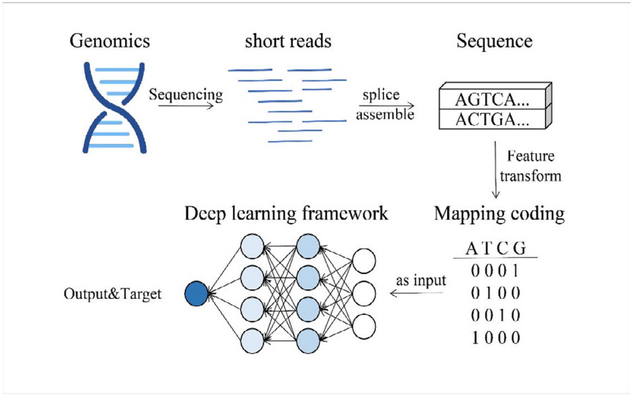
Electrical sequencing trace: deep learning to get sequence information



The system is designed to give ultra-high read length (hundreds of kb) and detect modified bases.
Large improvements in sequence quality thanks to base calling algorithms

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

Deep learning workflow in genome annotation



Brief Bioinform, Volume 25, Issue 3, May 2024, bbae138, https://doi.org/10.1093/bib/bbae138

#sanita2030

www.sanita2030.it

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

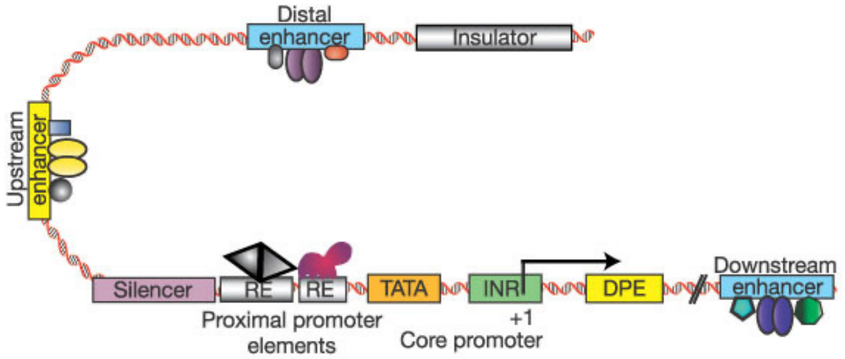
- Identification and classification of transposable elements
- Identification of protein-coding genes
  - Identification and prediction of splice sites
  - Alternative splicing events
- Identification of regulatory elements
  - Promoter recognition
  - Enhancer identification
  - Identification of transcription factor binding sites

#sanita2030

www.sanita2030.it

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

Cis-acting transcriptional control: a complex modular system



Levine & Tjian, Nature, 2003

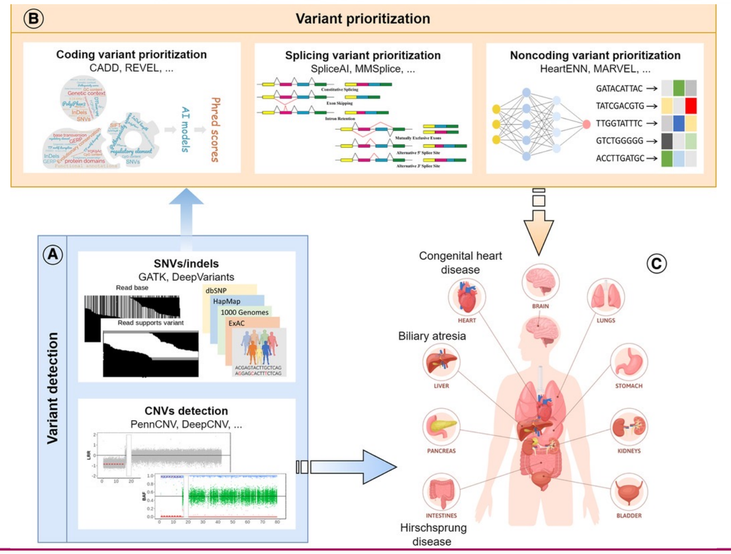
#sanita2030

www.sanita2030.it

L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

- DeepArk, a deep learning model for studying cis-regulatory activity across four extensively researched species, accurately predicts thousands of different regulatory features, including chromatin states, histone marks and transcription factors.
- Despite predicting thousands of regulatory features, DeepArk may still miss modeling certain regulatory features due to data limitations. For example, it may not cover TF binding in rare cell types.
- In addition, DeepArk's predictions may vary across different contexts, increasing the complexity of analysis, and this complexity can make analysis challenging but also generate novel hypotheses for mechanistic experiments.

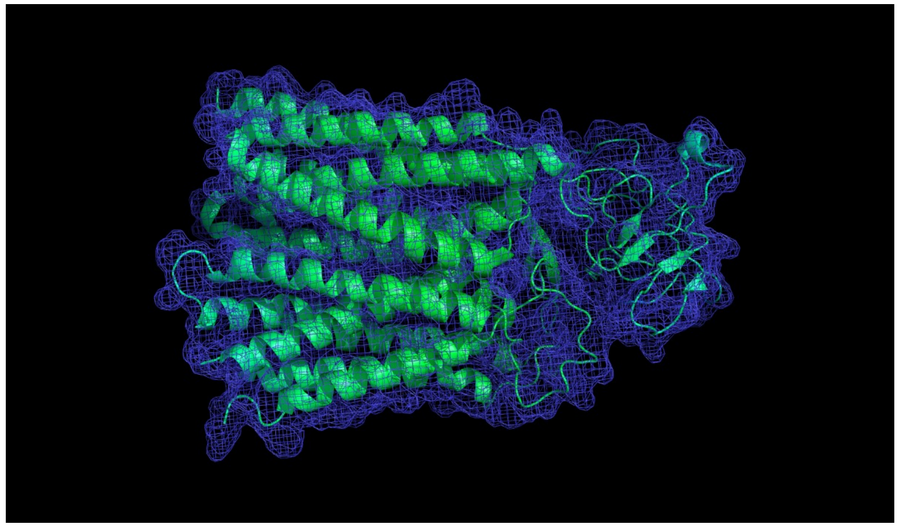
laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE



Lin et al.,Frontiers in Pediatrics, 2023

#sanita2030
www.sanita2030.it

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE



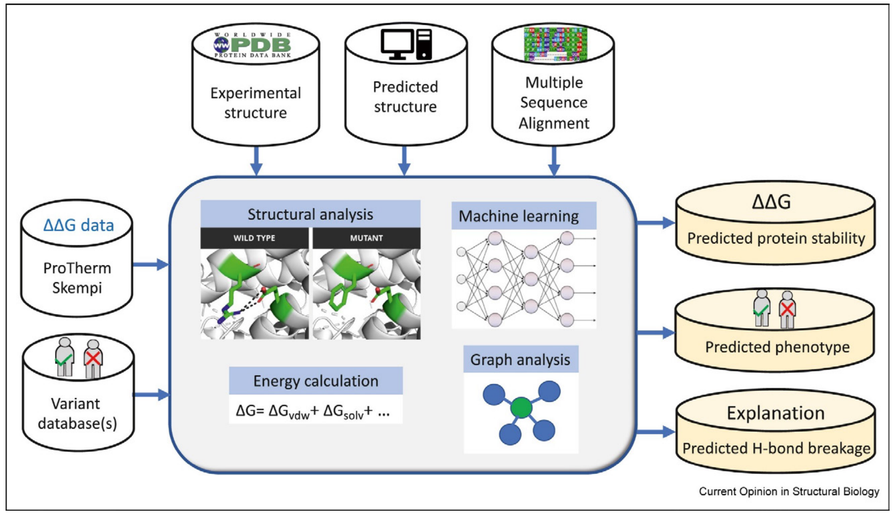
AlphaFold
RoseTTAFold
ESMFold

#sanita2030

Deep learning algorithms for protein structure prediction

www.sanita2030.it

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE



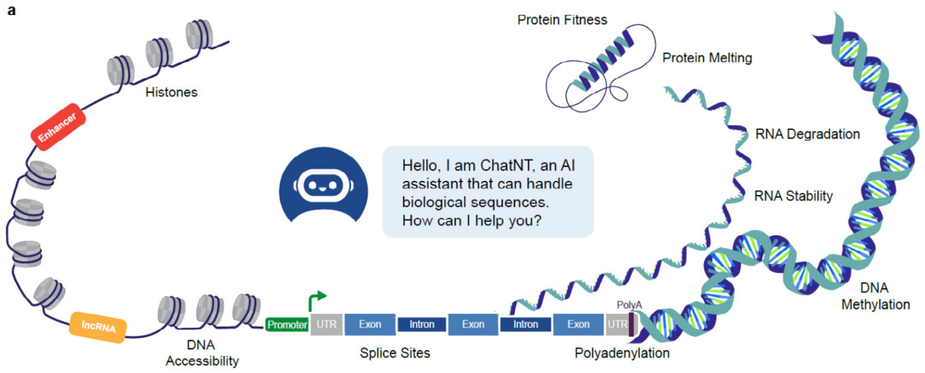
#sanita2030

Structure based variant effect predictor

www.sanita2030.it

laboratorio SANITÀ 2030
27-28 GIUGNO 2024 FRIULI VENEZIA GIULIA
VILLA MANIN CODROIPO - UDINE
L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

ChatNT: A Multimodal Conversational Agent for DNA, RNA and Protein Tasks



InstaDeep Ltd, London, UK, BioNTech, Mainz, Germany

Delitti in materia di violazione del diritto d'autore (Art. 25-novies, D.Lgs. n. 231/2001) [articolo aggiunto dalla L. n. 99/2009]

- Messa a disposizione del pubblico, in un sistema di reti telematiche, mediante connessioni di qualsiasi genere, di un'opera dell'ingegno protetta, o di parte di essa (art. 171, legge n.633/1941 comma 1 lett. a) bis)
- Reati di cui al punto precedente commessi su opere altrui non destinate alla pubblicazione qualora ne risulti offeso l'onore o la reputazione (art. 171, legge n.633/1941 comma 3)
- Abusiva duplicazione, per trarne profitto, di programmi per elaboratore; importazione, distribuzione, vendita o detenzione a scopo commerciale o imprenditoriale o concessione in locazione di programmi contenuti in supporti non contrassegnati dalla SIAE; predisposizione di mezzi per rimuovere o eludere i dispositivi di protezione di programmi per elaboratori (art. 171-bis legge n.633/1941 comma 1)
- Riproduzione, trasferimento su altro supporto, distribuzione, comunicazione, presentazione o dimostrazione in pubblico, del contenuto di una banca dati; estrazione o reimpiego della banca dati; distribuzione, vendita o concessione in locazione di banche di dati (art. 171-bis legge n.633/1941 comma 2)
- Abusiva duplicazione, riproduzione, trasmissione o diffusione in pubblico con qualsiasi procedimento, in tutto o in parte, di opere dell'ingegno destinate al circuito televisivo, cinematografico, della vendita o del noleggio di dischi, nastri o supporti analoghi o ogni altro supporto contenente fonogrammi o videogrammi di opere musicali, cinematografiche o audiovisive assimilate o sequenze di immagini in movimento; opere letterarie, drammatiche, scientifiche o didattiche, musicali o drammatico musicali, multimediali, anche se inserite in opere collettive o composite o banche dati; riproduzione, duplicazione, trasmissione o diffusione abusiva, vendita o commercio, cessione a qualsiasi titolo o importazione abusiva di oltre cinquanta copie o esemplari di opere tutelate dal diritto d'autore e da diritti connessi; immissione in un sistema di reti telematiche, mediante connessioni di qualsiasi genere, di un'opera dell'ingegno protetta dal diritto d'autore, o parte di essa (art. 171-ter legge n.633/1941)
- Mancata comunicazione alla SIAE dei dati di identificazione dei supporti non soggetti al contrassegno o falsa dichiarazione (art. 171-septies legge n.633/1941)
- Fraudolenta produzione, vendita, importazione, promozione, installazione, modifica, utilizzo per uso pubblico e privato di apparati o parti di apparati atti alla decodificazione di trasmissioni audiovisive ad accesso condizionato effettuate via etere, via satellite, via cavo, in forma sia analogica sia digitale (art. 171-octies legge n.633/1941).

Torna all'inizio