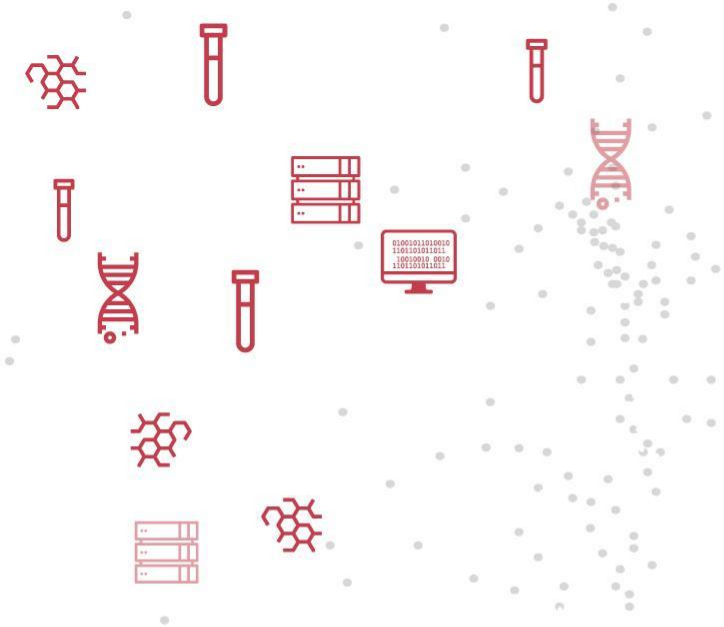
BITGENIA

TRANSCENDING GENOMICS

Impacto de la Inteligencia Artificial en Genómica Clínica

Bioq. NICOLÁS M. AGUIRRE
Analista de Datos Genómicos | Bitgenia
Contacto: nicolas.aguirre@bitgenia.com



Introducción

Inteligencia Artificial

Subdisciplina del campo de la informática que busca la creación de máquinas que puedan imitar comportamientos inteligentes.

- Vehículos Autónomos
- Generación de Imágenes
- Generación de Texto
- Imitación de voz

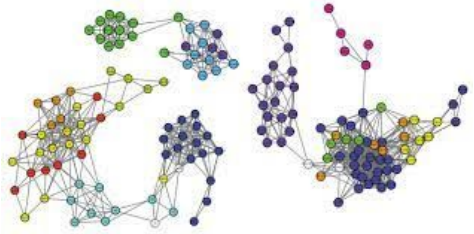
BITGENIA

Aprendizaje Automático (Machine Learning)

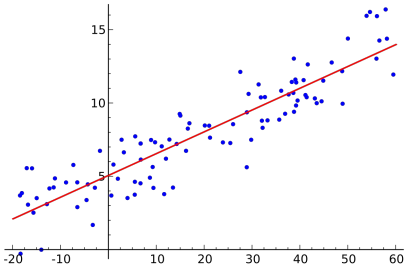
Rama del campo de la IA, que busca como dotar a las máquinas de aprendizaje.

Métodos:

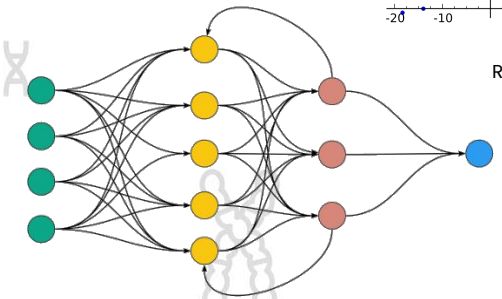
- Árboles de decisión
- Modelos de regresión
- Modelos de clasificación
- Técnicas de clusterización
- Redes Neuronales:
  - Información Jerarquizada
  - Aumento de capas de abstracción
  - Deep Learning



Clusterización

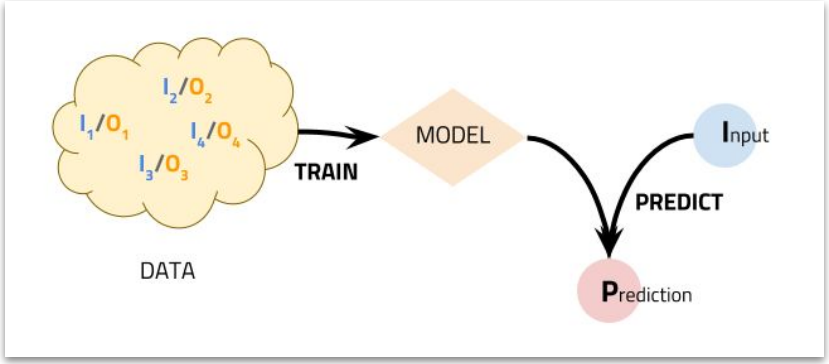


Regresión Lineal



Redes Neuronales

Cambio de Paradigma



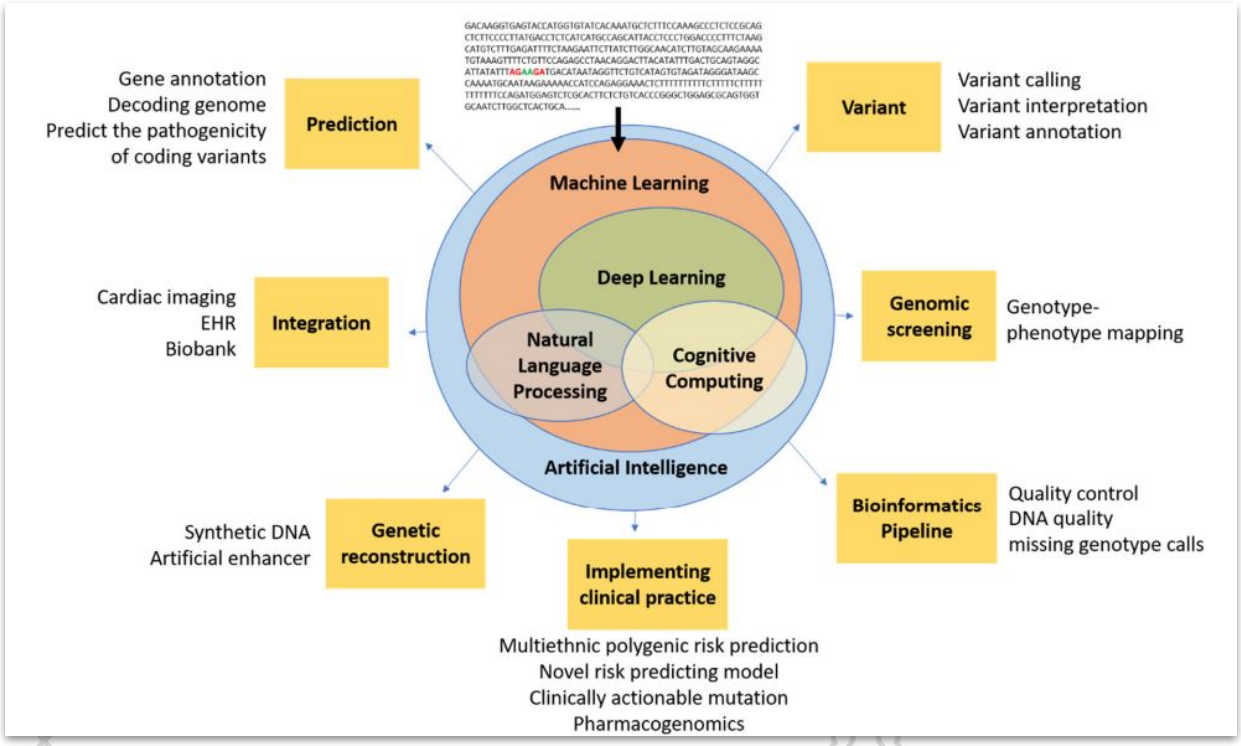
Anteriormente se programaba un algoritmo que seguía paso a paso una determinada tarea. **Había que contemplar cada variable posible**.

Con IA/ML se le da un objetivo y el programa va **aprendiendo hasta alcanzarlo**. Esto le otorga una capacidad disruptiva.

BITGENIA

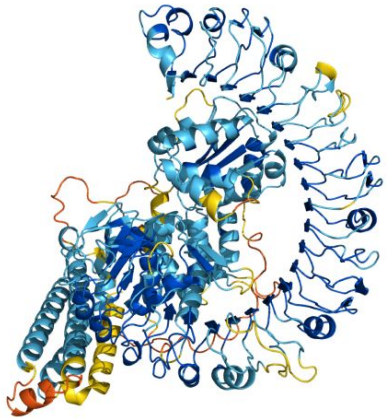
Actualidad

Inteligencia Artificial en Genómica

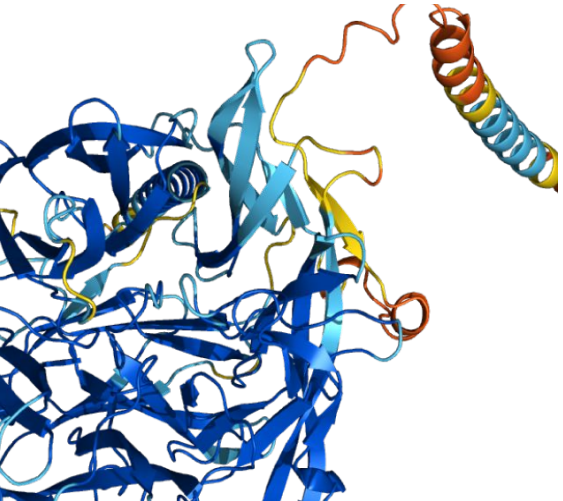


AlphaFold

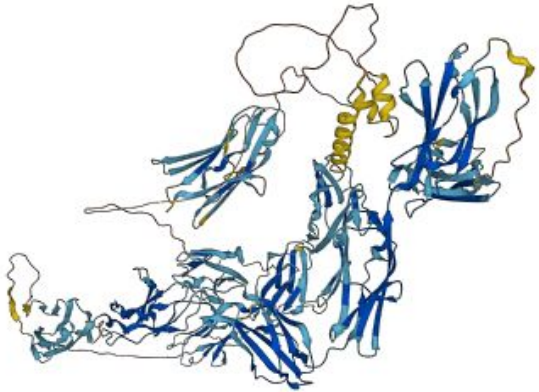
Es un Sistema de Inteligencia artificial que **predice** la estructura **3D** de las proteínas, **superando** en algunos casos la exactitud de los experimentos.



200 millones de predicciones, **50.000** estructuras proteicas humanas



Utiliza nuevas **redes neuronales** y entrenamientos basados en **restricciones evolutivas**, **físicas** y **geométricas**.

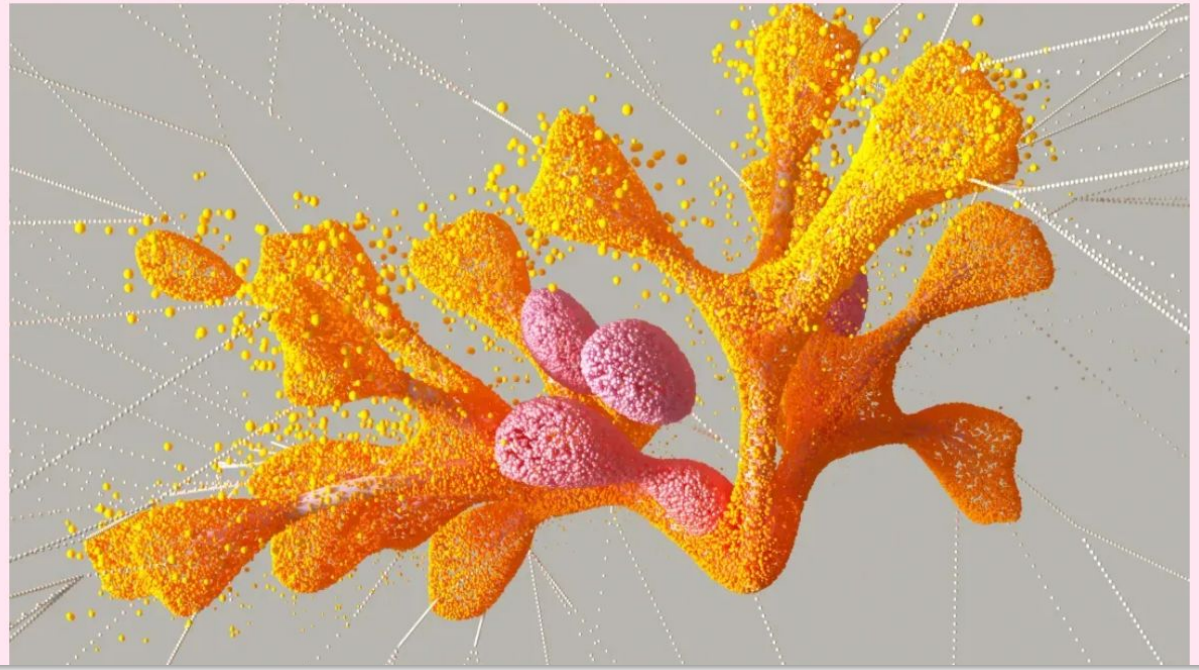


Alpha Missense

Research

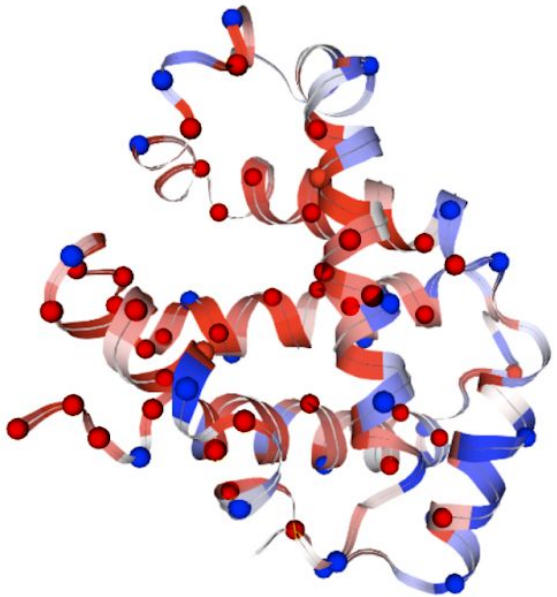
A catalogue of genetic mutations to help pinpoint the cause of diseases

September 19, 2023

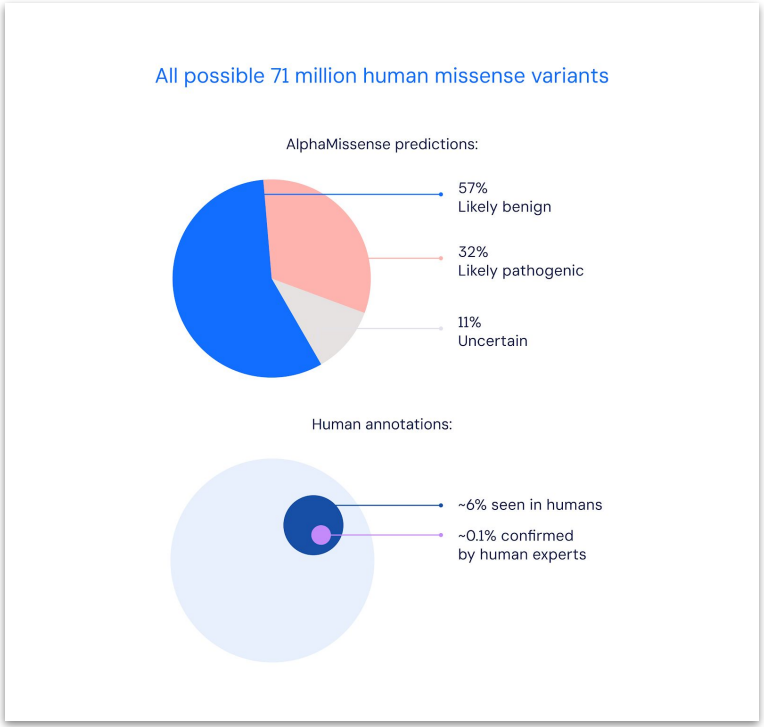


BITGENIA

Alpha Missense



Hemaglobin subunit beta (HBB)



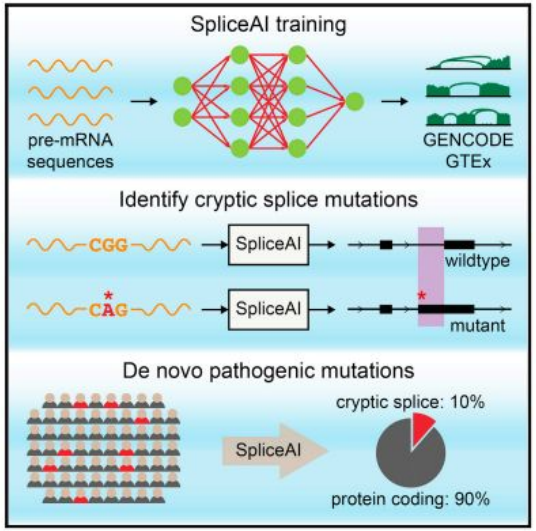
AlphaMissense predijo la patogenicidad de **71 millones de posibles variantes Missense**.

Clasificó el 89%, prediciendo el 57% como benignos y el 32% patógenos.

SpliceAI

Predicting Splicing from Primary Sequence with Deep Learning

Graphical Abstract



Authors

Kishore Jaganathan, Sofia Kyriazopoulou Panagiotopoulou, Jeremy F. McRae, ..., Serafim Batzoglou, Stephan J. Sanders, Kyle Kai-How Farh

Correspondence

kfarh@illumina.com

In Brief

A deep neural network precisely models mRNA splicing from a genomic sequence and accurately predicts noncoding cryptic splice mutations in patients with rare genetic diseases.

- Es una **red neuronal profunda** de 32 capas, que **predice el splicing** a partir de una secuencia de pre-ARNm
- **75%** de las variantes de splicing crípticas predichas **se confirman** en RNA-seq
- Las **variantes de splicing críptico** dan lugar con frecuencia a splicing alternativo, que puede derivar en **variantes truncantes**.

BITGENIA

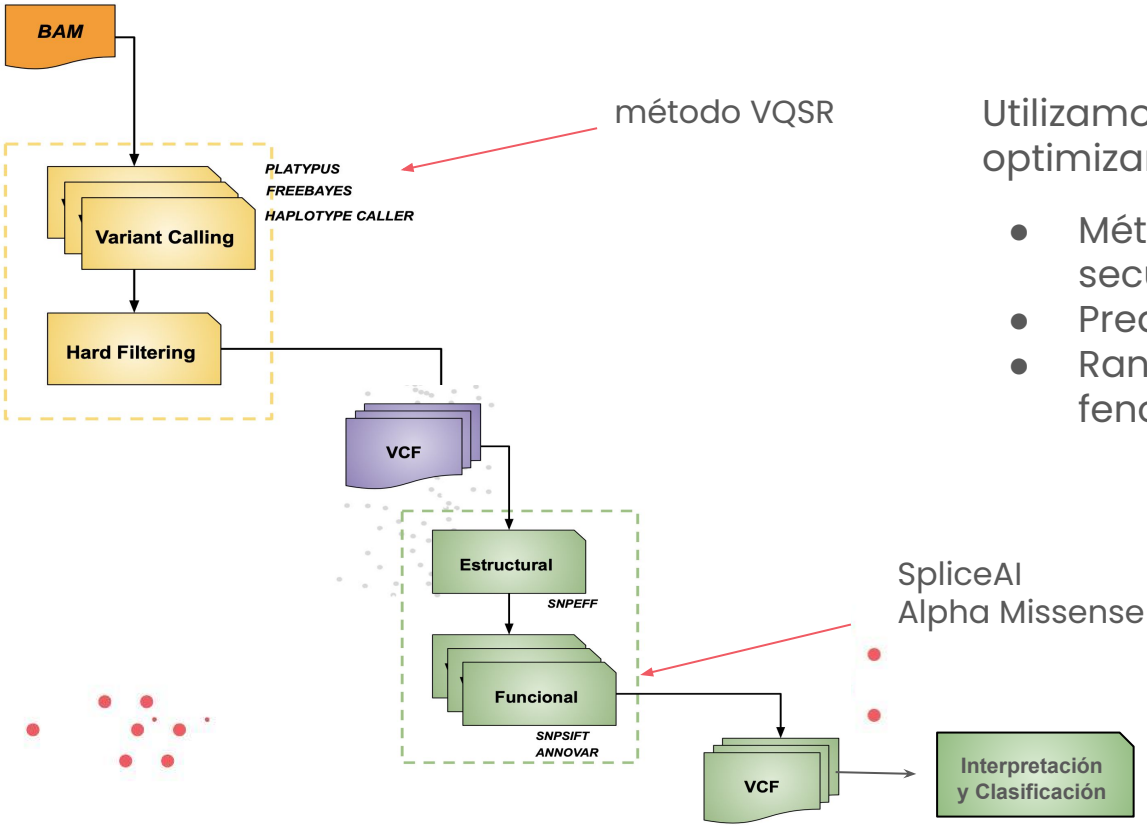
BITGENIA

TRANSCENDING GENOMICS

En Bitgenia

Venimos desarrollando e implementando herramientas que faciliten y mejoren el análisis de variantes genómicas

Incorporación de IA en Análisis Genómico



Utilizamos herramientas de IA para optimizar nuestros resultados.

- Métodos de control de calidad de la secuenciación
- Predicción computacional
- Ranking de de genes de acuerdo al fenotipo en estudio

Modelado bayesiano de la clasificación de variantes

Modeling the ACMG/AMP Variant Classification Guidelines as a Bayesian Classification Framework

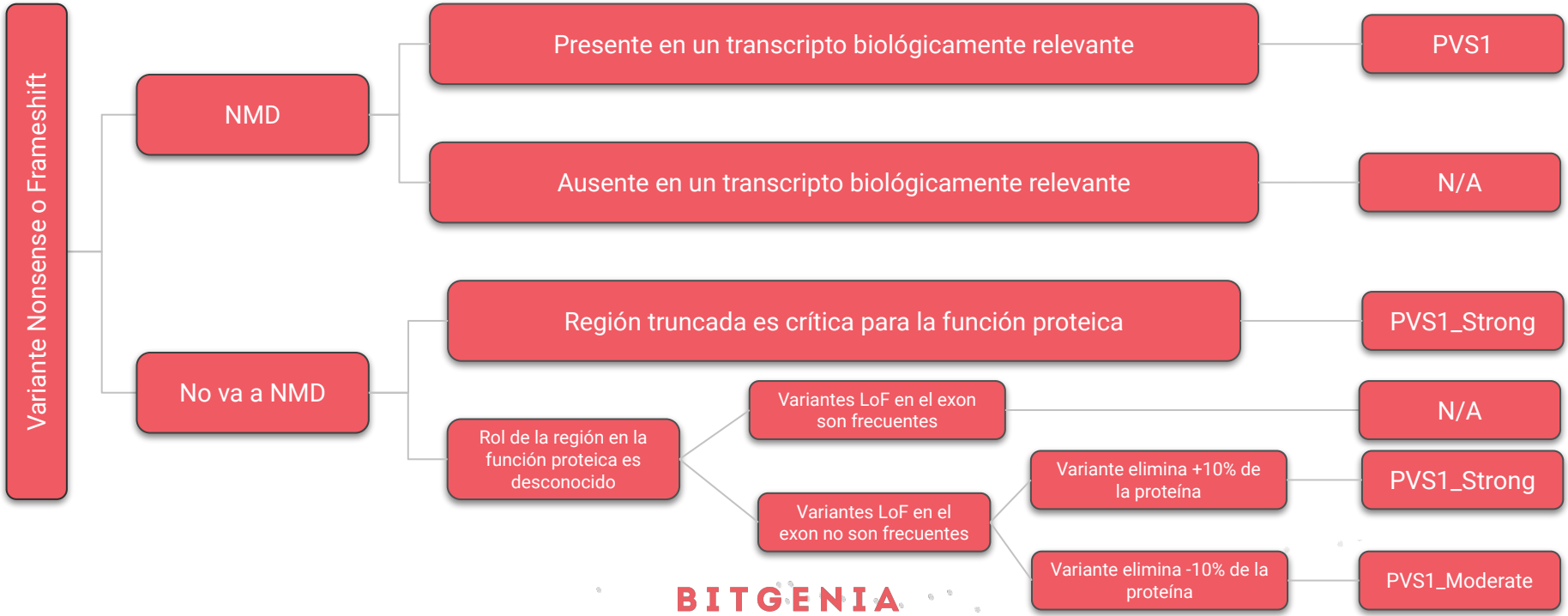
$$Post\_P = \frac{OddsPath * Prior\_P}{((OddsPath - 1) * Prior\_P + 1)}$$

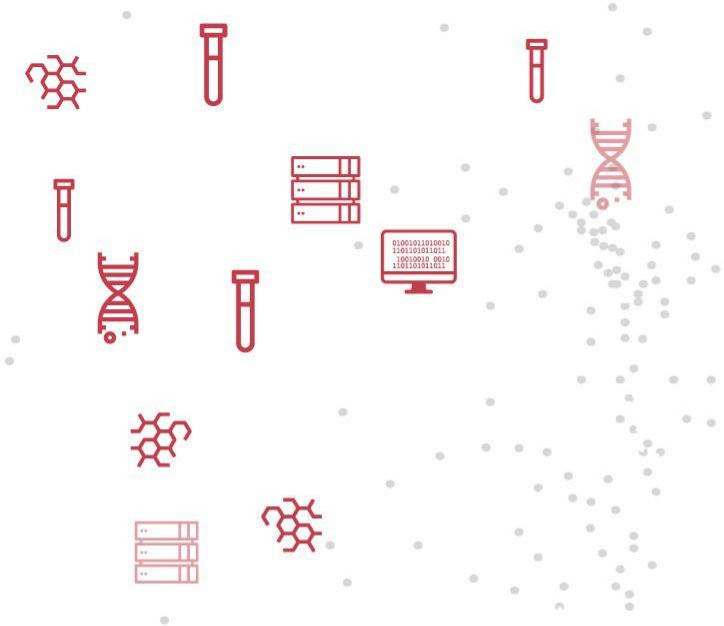
Sean V. Tavtigian[1,†], Marc S. Greenblatt[2], Steven M. Harrison[3], Robert L. Nussbaum[4], Snehit A. Prabhu[5], Kenneth M. Boucher[6], and Leslie G. Biesecker[7] on behalf of the ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI)[¥]

Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines

Sean V. Tavtigian,[1,2] Steven M. Harrison,[3] Kenneth M. Boucher,[2,4] and Leslie G. Biesecker[5]

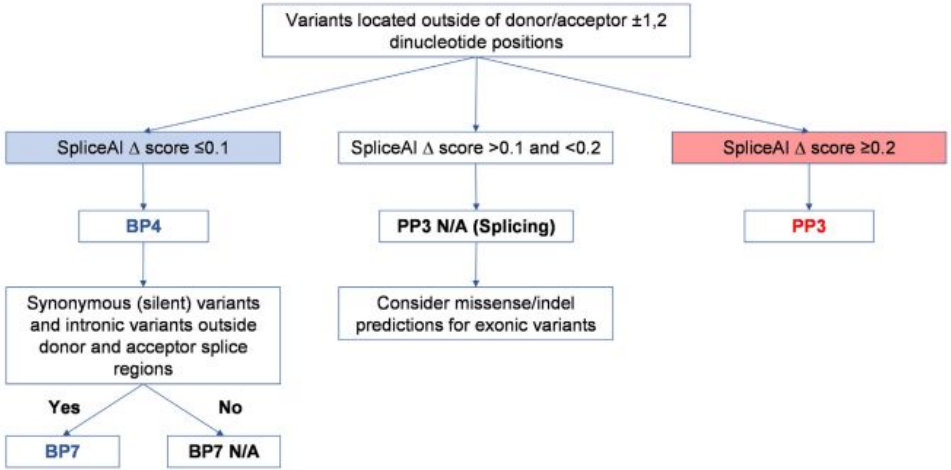
PVS1 (Criterio para variantes LoF)

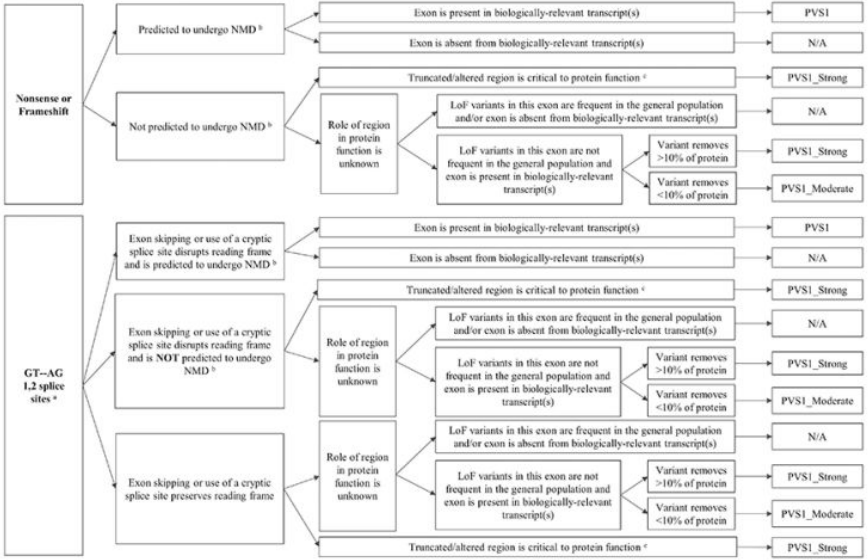




Futuro

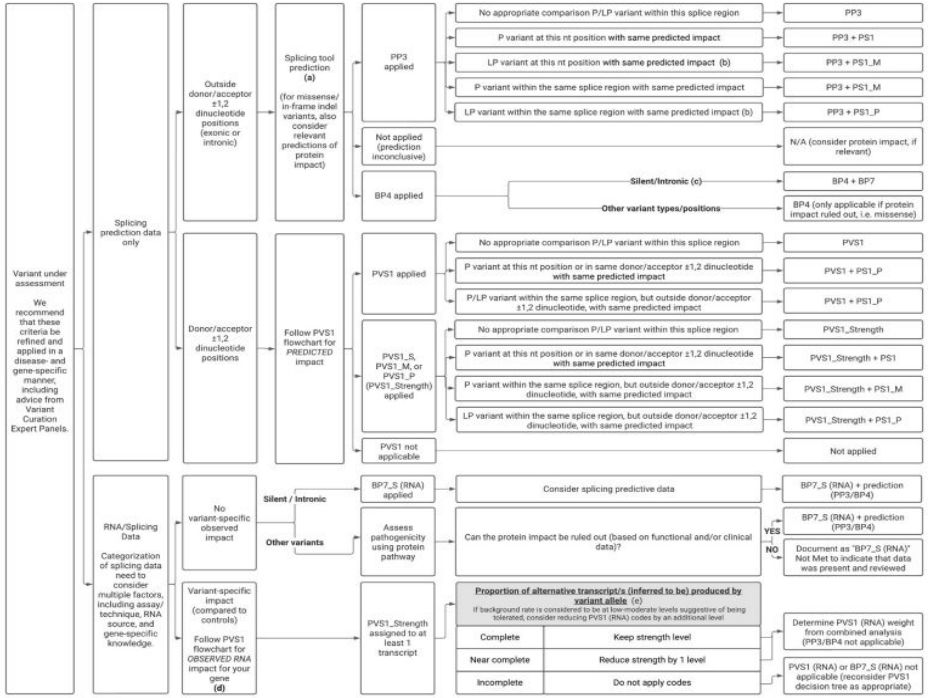
Árboles de Decisión

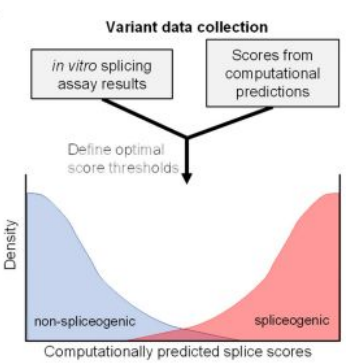




Using the ACMG/AMP framework to capture evidence related to predicted and observed impact on splicing: Recommendations from the ClinGen SVI Splicing Subgroup

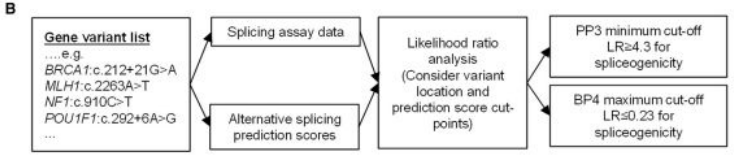
Logan C. Walker,[1] Miguel de la Hoya,[2] George A.R. Wiggins,[1] Amanda Lindy,[3] Lisa M. Vincent,[4] Michael T. Parsons,[5] Daffodil M. Canson,[5] Dana Bis-Brewer,[3] Ashley Cass,[6] Alexander Tchourbanov,[6] Heather Zimmermann,[6] Alicia B. Byrne,[7] Tina Pesaran,[6] Rachid Karam,[6] Steven M. Harrison,[6,7,*] Amanda B. Spurdle,[5,8] and ClinGen Sequence Variant Interpretation Working Group



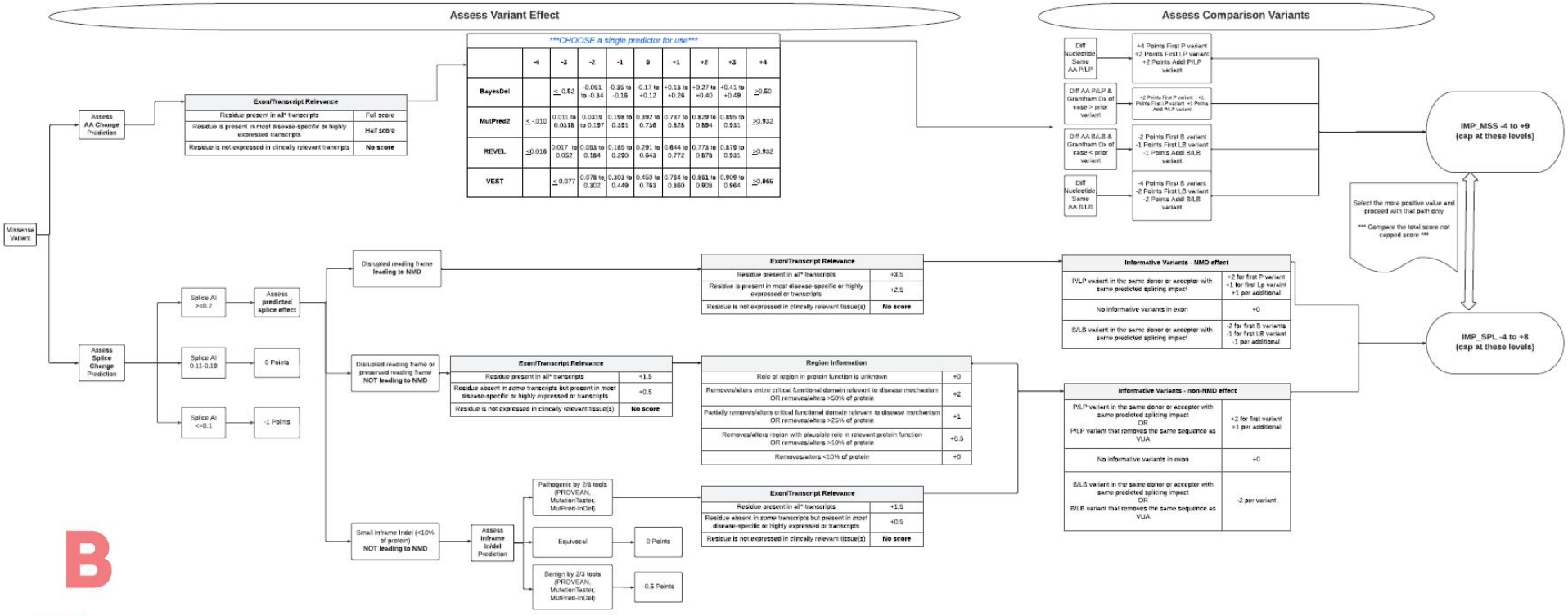


	Non-spliceogenic	Proportion	Spliceogenic	Proportion	Likelihood of event
Low prediction score	n1	P1= n1/(n1+n2+n3)	n4	P4= n4/(n4+n5+n6)	LR1 = P4/P1
Intermediate	n2	P2= n2/(n1+n2+n3)	n5	P5= n5/(n4+n5+n6)	LR2 = P5/P1
High prediction score	n3	P3= n3/(n1+n2+n3)	n6	P6= n6/(n4+n5+n6)	LR3 = P6/P3

Likelihood ratio for spliceogenic: 350 (Very strong), 17.7 (Strong), 4.3 (Moderate), 2.08 (Supporting).
Likelihood ratio for non-spliceogenic: 0.481 (Supporting), 0.233 (Moderate), 0.053 (Strong), 0.003 (Very strong)

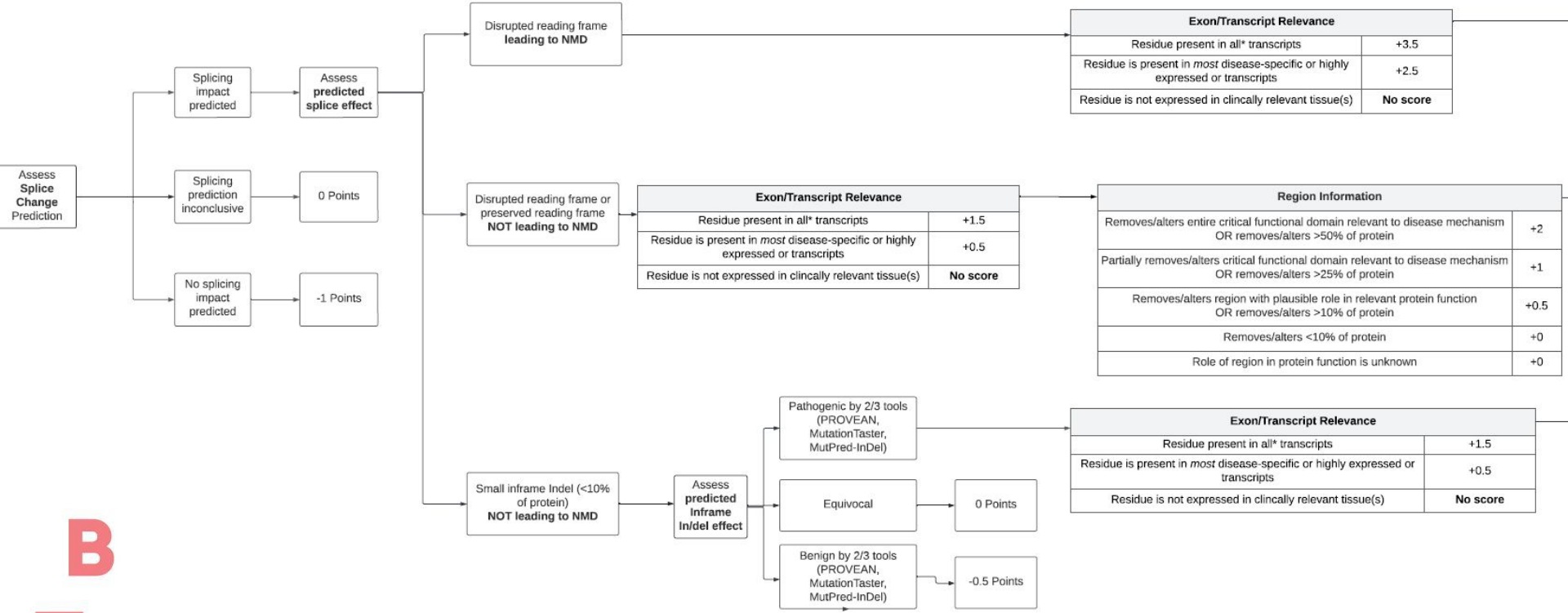


Proximos avances



B

Proximos avances



Un vistazo al futuro del análisis genómico

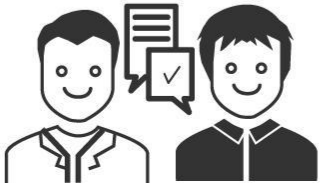
BITGENIA

Convergencia de la información en árboles de decisión que luego se transformarán en **redes neuronales** permitiendo obtener clasificaciones a partir de **múltiples parámetros biológicos involucrados**.

Se están construyendo múltiples bases de datos que incluyen **todas las ómicas**.

Estamos ante la era de la **información**.

Conclusiones



Conclusión

La IA/ML está aportando un **nuevo modo de explotación** de la información brindándonos la capacidad de hacer análisis **n-dimensionales**.

La información está en proceso de revisión para **eliminar los sesgos humanos** ya que estos van a estar incluidos en cualquier modelo de aprendizaje automatizado.

La potencia de la inteligencia artificial está a la vista, es nuestro deber formar parte de los grupos de decisión para hacer un **uso ético de los mismos**.

Referencias

¿Qué es el Machine Learning?¿Y Deep Learning? Un mapa conceptual | DotCSV

Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2

Tunyasuvunakool, K., Adler, J., Wu, Z. et al. Highly accurate protein structure prediction for the human proteome. Nature 596, 590–596 (2021). https://doi.org/10.1038/s41586-021-03828-1

Overview of ACMG/AMP v4 Sequence Variant Guidelines - ClinGen | Clinical Genome Resource

Tavtigian SV, Greenblatt MS, Harrison SM, Nussbaum RL, Prabhu SA, Boucher KM, Biesecker LG; ClinGen Sequence Variant Interpretation Working Group (ClinGen SVI). Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. Genet Med. 2018 Sep;20(9):1054-1060. doi: 10.1038/gim.2017.210. Epub 2018 Jan 4. PMID: 29300386; PMCID: PMC6336098.

Tavtigian SV, Harrison SM, Boucher KM, Biesecker LG. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. Hum Mutat. 2020 Oct;41(10):1734-1737. doi: 10.1002/humu.24088. Epub 2020 Aug 30. PMID: 32720330; PMCID: PMC8011844.

Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, Kosmicki JA, Arbelaez J, Cui W, Schwartz GB, Chow ED, Kanterakis E, Gao H, Kia A, Batzoglou S, Sanders SJ, Farh KK. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019 Jan 24;176(3):535-548.e24. doi: 10.1016/j.cell.2018.12.015. Epub 2019 Jan 17. PMID: 30661751.

https://alphafold.ebi.ac.uk/

https://www.deepmind.com/blog/alphamissense-catalogue-of-genetic-mutations-to-help-pinpoint-the-cause-of-diseases

BITGENIA

BITGENIA

TRANSCENDING
GENOMICS

Bioq. NICOLÁS M. AGUIRRE
Analista de Datos Genómicos | Bitgenia
Contacto: nicolas.aguirre@bitgenia.com

INFO@BITGENIA.COM - 54 9 11 6724 2109 - BUENOS AIRES, ARGENTINA

WWW.BITGENIA.COM