

CSVS: El Servidor Colaborativo de Variabilidad Genética Española

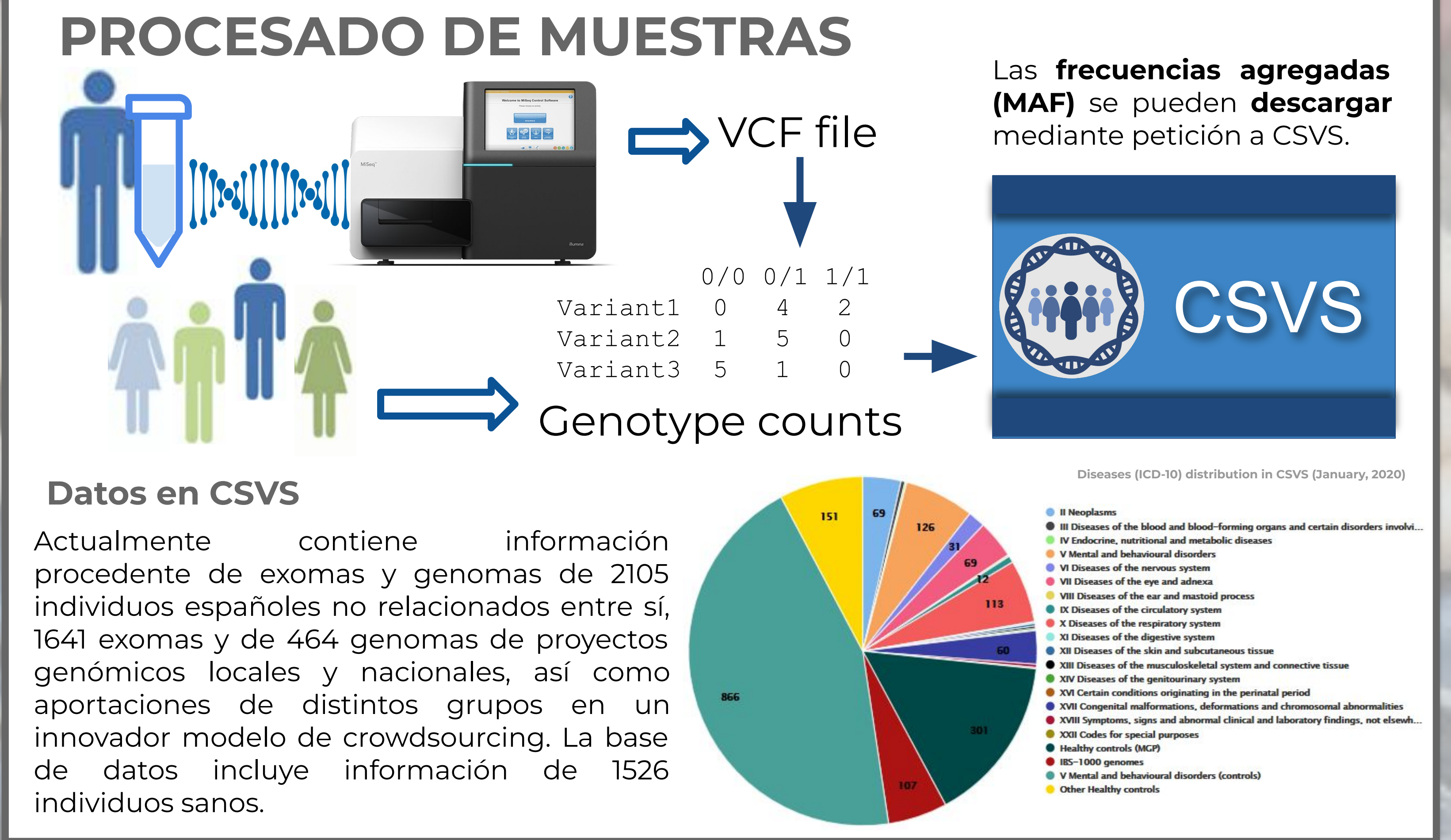
Maria Peña-Chilet^{1,2,3}, Gema Roldán¹, Jose L. Fernández-Rueda¹, Gerrit Bostlema¹, Virginia Aquino¹, Daniel Lopez-Lopez^{1,2}, Francisco M. Ortuño^{1,2,4}, Javier Pérez-Florido^{1,2,5}, Anna Gonzalez-Neira⁶, Rosario Carmona¹, Joaquín Dopazo^{1,2,3,7}

¹Plataforma Andaluza de Medicina Computacional, Fundación Progreso y Salud. Sevilla. ²Grupo de Medicina computacional y de Sistemas, Instituto de Biomedicina de Sevilla (IBIS). Sevilla. ³Centro de Investigaciones Biomédicas en Enfermedades Raras (CIBERER). ⁴Departamento de Ingeniería de Computadores, Automática y Robótica. Universidad de Granada. ⁵Hospital Virgen del Rocío. Sistema Andaluz de Salud. Sevilla. ⁶Centro Nacional de investigaciones oncológicas (CNIO). Madrid. ⁷Instituto Nacional de Bioinformática, INB-ELIXIR-es.

INTRODUCCIÓN

El conocimiento de la variabilidad genética local es de gran importancia en la medicina personalizada, ya que permite un análisis más preciso, siendo un factor crítico para el descubrimiento de nuevos genes y variantes genéticas de enfermedad.

Los resultados obtenidos de proyectos genómicos sugieren que el conocimiento de la variabilidad genética es de enorme importancia en medicina personalizada y, además, es un factor crítico para el descubrimiento de nuevas variantes de enfermedad, sin embargo, las frecuencias alélicas difieren considerablemente entre poblaciones, por esto es importante contar con un catálogo local de la variabilidad genómica, ya que permite descartar aquellas variantes candidatas por baja frecuencia en otras poblaciones, pero que tienen gran representación en nuestra población.



ANÁLISIS DE LA VARIABILIDAD ESPAÑOLA

CSVS (Collaborative Spanish Variant Server) proporciona información sobre la variabilidad genómica de la población española, muy útil para filtrar polimorfismos y variaciones locales durante el proceso de priorización de genes candidatos a explicar una enfermedad. Actualmente contiene información procedente de exomas y genomas de 2105 individuos españoles no relacionados entre sí, que incluyen tanto individuos sanos como afectados de alguna patología. Además, cuenta con los siguientes recursos:

SATURATION POLYGENIC RISK SCORE PHARMACOGENOMICS SECONDARY FINDINGS

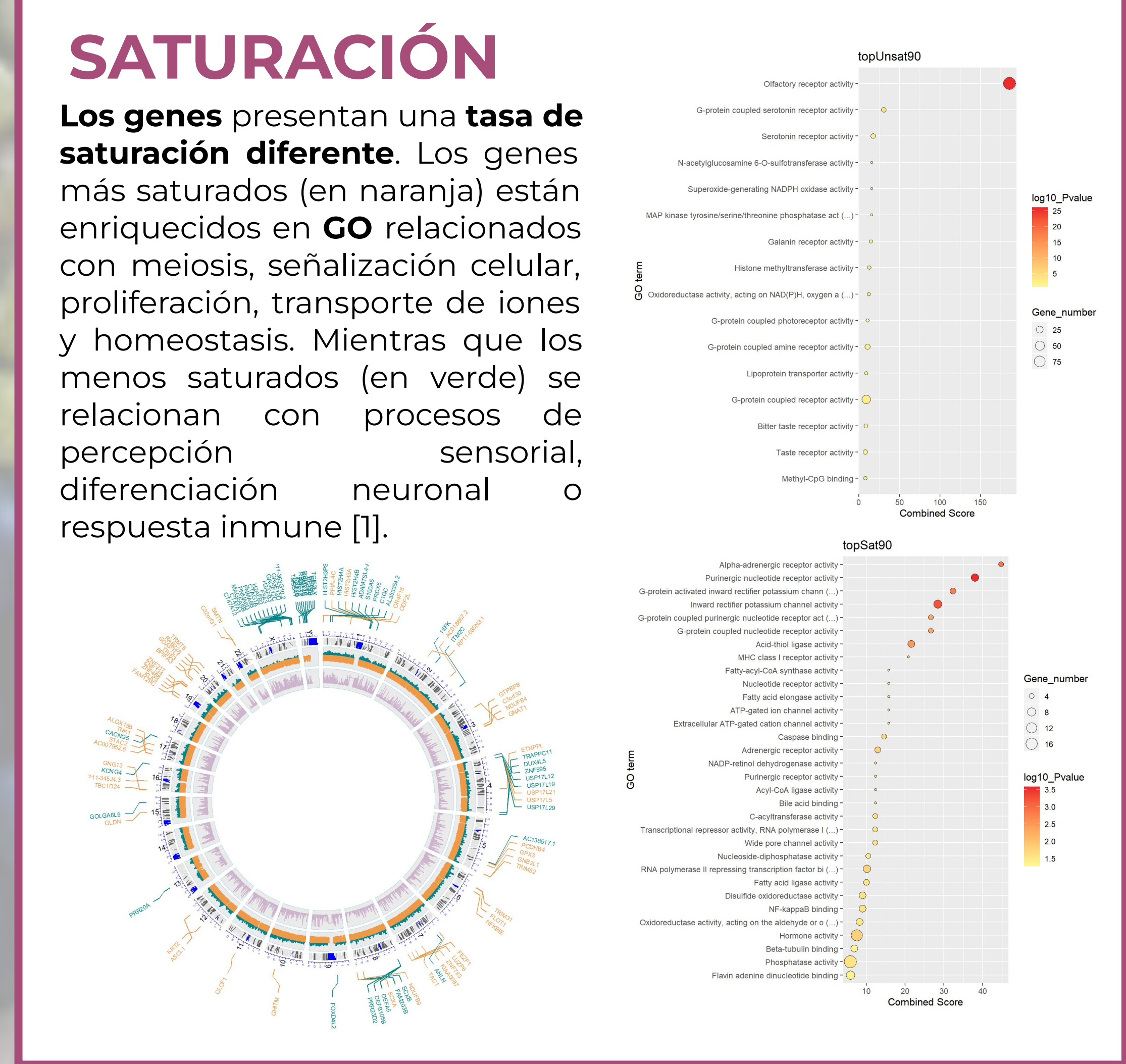
CONTROLES DE CALIDAD

Para garantizar la **fiabilidad de los datos incluidos en CSVS** implementamos **controles de identidad** que no comprometen la **privacidad** ni posibilitan la **reidentificación de individuos**. [1]

- **Control de artefactos y duplicaciones/parentesco:** modelo estadístico a partir de la distribución normal del **ratio de variantes privadas por muestra** para excluir: (i) muestras con alto ratio de variantes privadas (artefactos, *upper outlier*); (ii) muestras con bajo ratio de variantes privadas (duplicidades o parentescos directos, *lower outlier*).
- **Control poblacional:** detección de individuos no españoles mediante **machine learning** entrenado con biomarcadores propios de la **estructura poblacional de los 1000 genomas**.

HALLAZGOS SECUNDARIOS

CSVS anota las variantes clasificadas como posible hallazgo secundario según los criterios establecidos por ACMG para una lista de 78 genes accionables (ACMG V3.1). Proporciona además el fenotipo asociado, el término OMIM y el modo de herencia en cada caso. [2]



VARIABILIDAD FARMACOGENÓMICA

CSVS muestra a nivel de cada variante los haplotipos en los que puede estar involucrada, con enlaces a la base de datos PHARMGKB (<https://www.pharmgkb.org/>). Esta información resulta útil para conocer las variantes genéticas más prevalentes en población española que condicionan la respuesta a fármacos [3].

PRS

(PRS, Polygenic Risk Scores) Proporcionan una medida cuantitativa personalizada de la predisposición genética a una enfermedad, y de la contribución genética a rasgos cuantitativos. CSVS permite consultar y visualizar la distribución de las estimas de riesgo poligénico para todos los PRS del catálogo público <https://www.pgscatalog.org/>.

[1] Peña-Chilet et al.; Spanish Exome Crowdsourcing Consortium; Carracedo Á, Alonso Á, Dopazo J. CSVS, a crowdsourcing database of the Spanish population genetic variability. Nucleic Acids Res. 2021 Jan 8;49(D1):D1130-D1137. doi: 10.1093/nar/gkaa794. PMID: 32990755; PMCID: PMC7778906.
 [2] Miller DT, et al.; ACMG Secondary Findings Working Group. Electronic address: documents@acmg.net. ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2022 Jul;24(7):1407-1414. doi: 10.1016/j.gim.2022.04.006. Epub 2022 Jun 17. PMID: 35802134.
 [3] Nunez-Torres R, et al. A Comprehensive Analysis of 21 Actionable Pharmacogenes in the Spanish Population: From Genetic Characterisation to Clinical Impact. Pharmaceutics. 2023 Apr 19;15(4):1286. doi: 10.3390/pharmaceutics15041286. PMID: 37111771; PMCID: PMC10140932.