

J. Mendoza Fernández. ⁽¹⁾, K. Dongo Pflucker⁽¹⁾; S. Dávila Paico⁽¹⁾; J. Cisneros Tinco⁽¹⁾, M. Acosta Aliaga⁽¹⁾; S. Díaz Montes⁽¹⁾; B. Granda Arana⁽¹⁾; P. Mora Alferez⁽¹⁾, Y. Sulcahuamán Allende⁽¹⁾

(1) Instituto Nacional de Enfermedades Neoplásicas

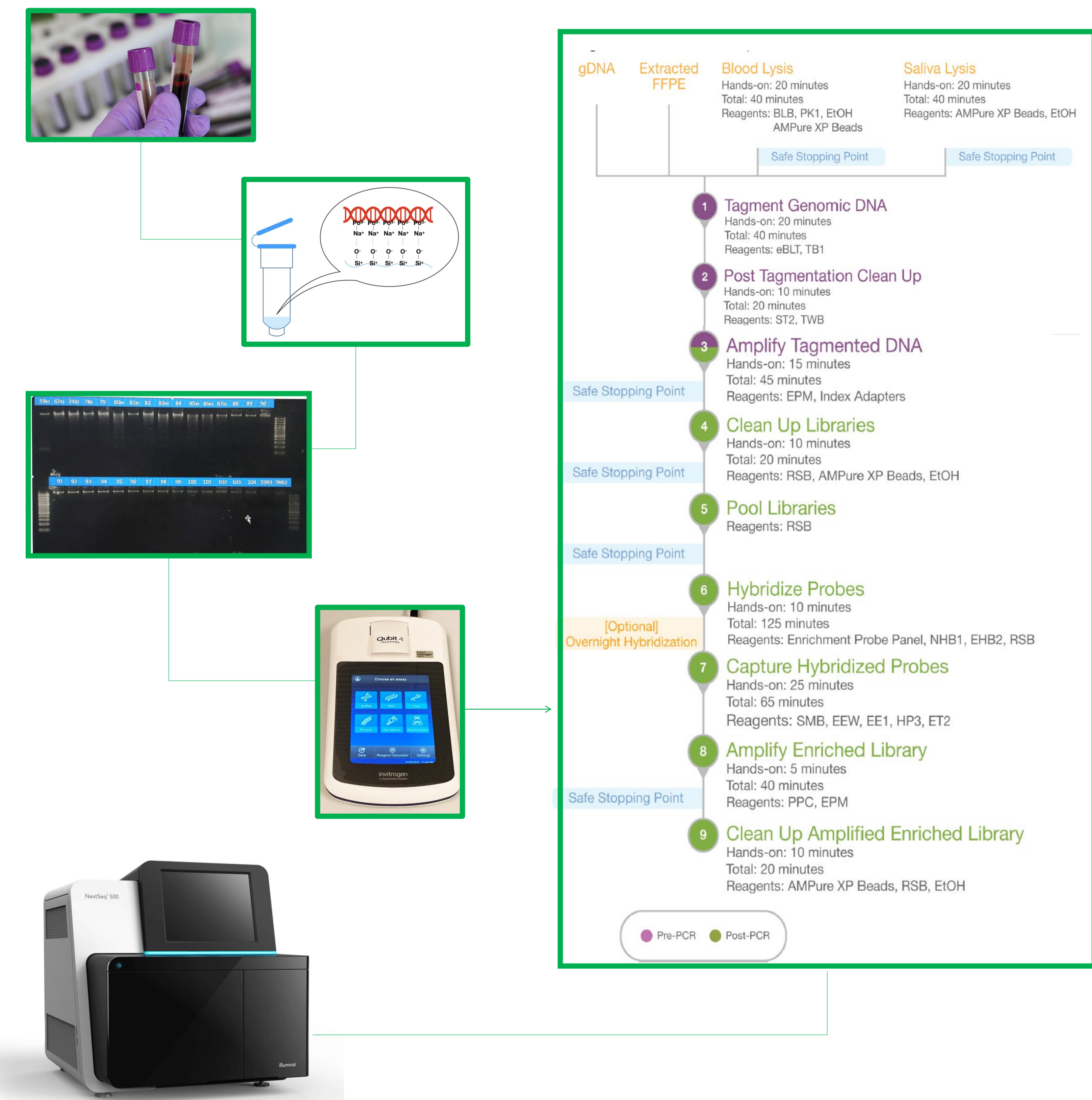
Introducción

En el laboratorio de Biología Molecular del Equipo Funcional de Genética y Biología Molecular del INEN se ha implementado exitosamente la tecnología NGS de *Illumina*. Un aspecto crucial en el análisis de los datos generados por la secuenciación masiva es la interpretación de las variantes genéticas identificadas. En este sentido, la *American College of Medical Genetics and Genomics* (ACMG) ha desarrollado una guía para la interpretación de variantes genéticas, brindando criterios detallados para evaluar la patogenicidad de las mismas (frecuencia en la población, uso de herramientas *in silico* para predecir el impacto funcional, registro de estudios funcionales y otras líneas de evidencia). La aplicación de estos criterios de clasificación ayuda a los profesionales de la salud, pacientes y familiares a tomar decisiones informadas sobre el diagnóstico y pronóstico en el marco del asesoramiento genético adecuado.

En este póster, presentaremos los avances logrados en la implementación de un flujo de trabajo multidisciplinario, biólogo molecular – médico genetista, resaltando finalmente la importancia de los hallazgos genéticos relevantes para el paciente y su entorno familiar. Esta *Data* enriquecedora del panorama genético de la población peruana constituye un respaldo valioso para la toma de decisiones en salud pública.

Metodología

El flujo de trabajo inicia con la consulta genética, asesoría genética y luego se procede con la toma de muestra para la extracción de ADN genómico, autorizada mediante un consentimiento informado. Se evalúa la calidad del ácido nucleico extraído por medio de electroforesis (agarosa al 3%), así como por la concentración determinada por fluorometría. Se requiere ADNg íntegro y >50ng totales como *input*. Finalmente el kit utilizado es *Illumina DNA Prep with Enrichment* para la preparación de librerías.



Los archivos FASTQ obtenidos para cada muestra son descargados de la nube de *Illumina* (*BaseSpace*) y subidos a la plataforma de *SOPHIA GENETICS*. Luego de aplicar el pipeline para el panel se evalúan los parámetros de calidad.

Table 1.1: Overview of datasets.

Sample ID	Number of reads	Number of Mapped reads	Number of reads with problematic mate	Percentage of Mapped reads
200105769-776-S1	3,373,716	3,346,537	44,383	99.19%
200105760-777-S2	4,493,158	4,463,084	68,550	99.33%
200105761-778-S3	4,811,448	4,781,902	80,840	99.39%
200105762-779-S5	5,345,298	5,312,486	65,273	99.39%
200105763-780-S6	4,666,538	4,628,457	57,485	99.18%
200105764-781-S7	4,711,422	4,682,773	76,767	99.39%
200105765-782-S8	3,793,304	3,770,449	53,131	99.40%
200105766-783-S9	3,928,134	3,900,351	47,851	99.29%
200105767-784-S10	4,604,654	4,573,897	63,513	99.33%
200105768-785-S11	4,146,452	4,117,310	52,894	99.30%
200105769-786-S12	4,045,352	4,018,621	41,923	99.36%
200105770-787-S13	3,702,656	3,681,849	40,333	99.44%
200105771-788-S14	3,038,850	3,010,277	24,791	99.22%
200105772-789-S15	3,508,820	3,485,685	41,735	99.34%
200105773-790-S16	2,291,670	2,279,067	12,675	99.45%
200105774-791-S17	2,723,372	2,708,053	30,687	99.44%
200105775-792-S18	2,778,802	2,763,822	30,468	99.43%
200105776-793-S19	2,549,164	2,534,467	26,447	99.42%
200105777-794-S20	3,081,948	3,056,482	26,882	99.17%
200105778-795-S21	3,154,162	3,136,865	37,707	99.45%

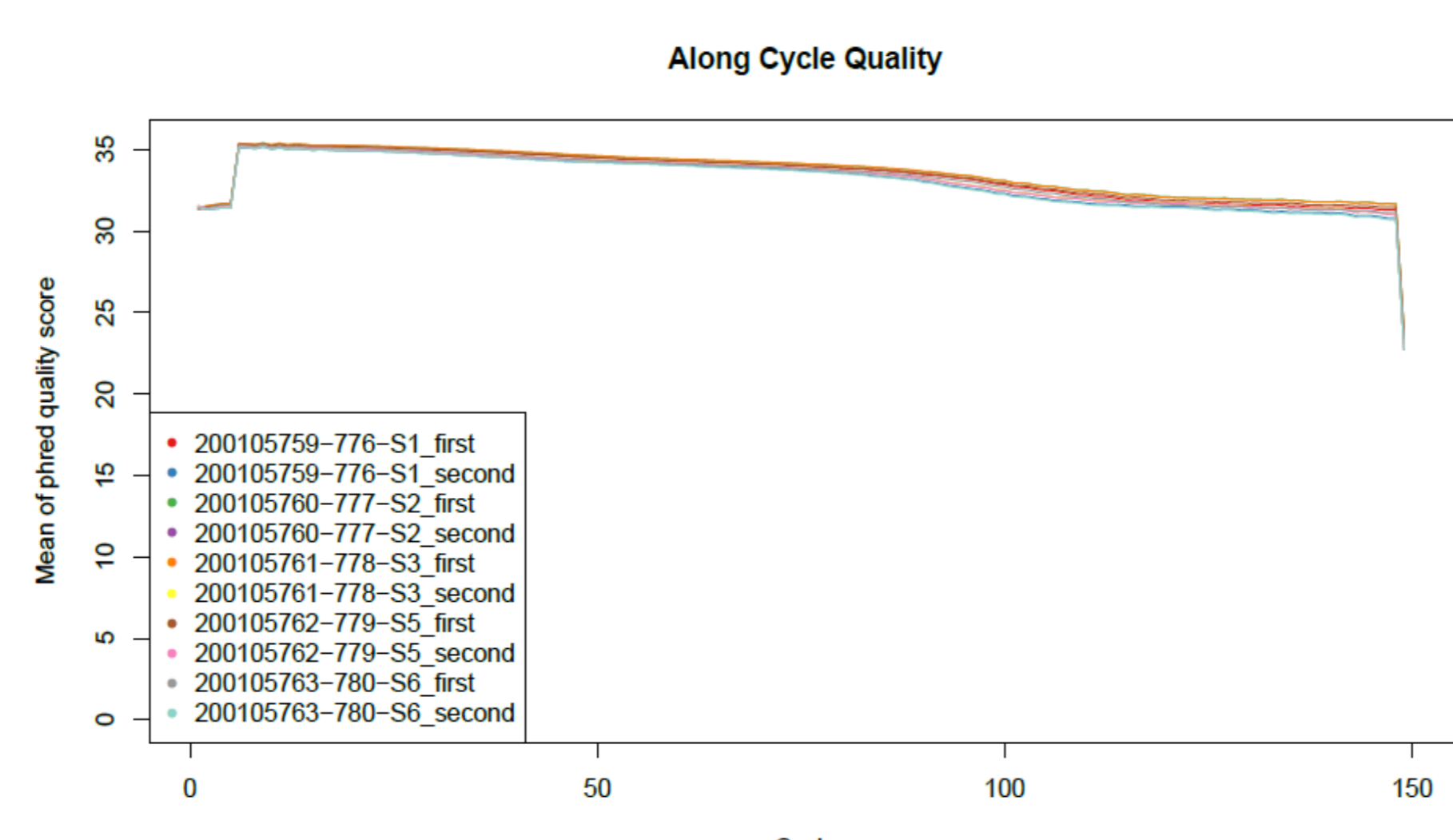
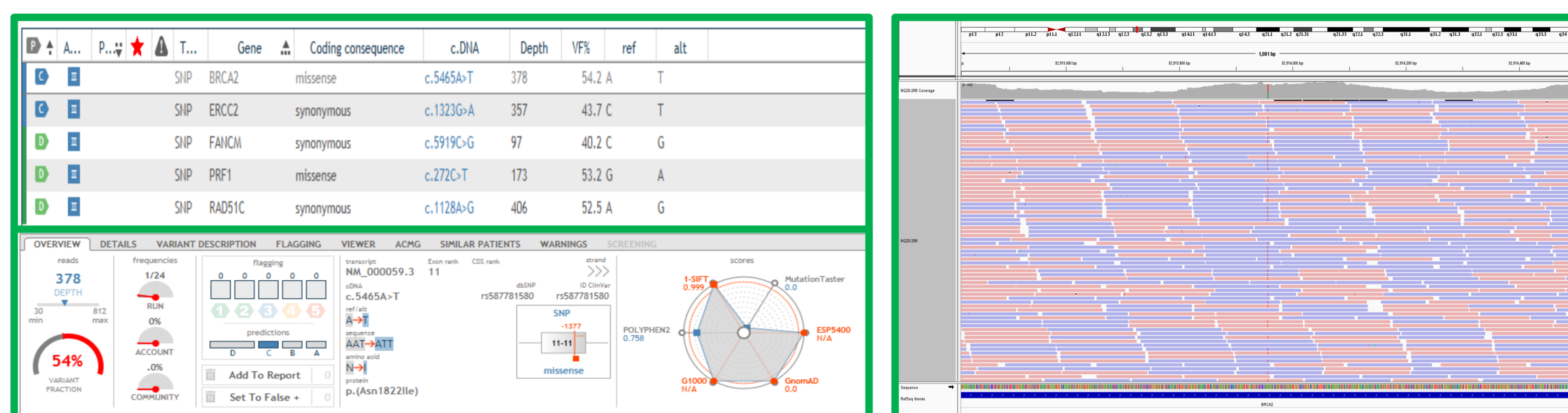


Table 1.6: Percentage of target regions with certain coverage.

Sample ID	25x	50x	100x	200x	500x	1000x	Coverage 10% quantile	Coverage Heterogeneity
200105769-776-S1	99.01%	98.57%	97.74%	96.29%	84.02%	11.45%	4374	2.89%
200105760-777-S2	99.09%	98.85%	98.15%	97.04%	92.44%	49.08%	571x	2.94%
200105761-778-S3	99.08%	98.80%	98.20%	97.26%	93.97%	61.58%	650x	2.90%
200105762-779-S5	99.19%	98.84%	98.39%	97.52%	94.33%	75.70%	695x	2.82%
200105763-780-S6	99.16%	98.85%	98.19%	97.21%	92.72%	51.33%	578x	2.81%
200105764-781-S7	99.00%	98.77%	98.10%	97.41%	93.17%	56.26%	600x	3.07%
200105765-782-S8	99.00%	98.68%	97.89%	96.81%	88.78%	26.61%	477x	2.97%
200105766-783-S9	99.02%	98.72%	97.92%	96.79%	89.92%	29.92%	493x	2.81%
200105767-784-S10	99.13%	98.88%	98.14%	97.16%	92.87%	52.11%	588x	2.87%
200105768-785-S11	99.07%	98.76%	98.02%	96.83%	91.99%	38.39%	522x	2.92%
200105769-786-S12	99.04%	98.73%	97.96%	96.76%	90.16%	32.85%	504x	2.92%
200105770-787-S13	98.47%	97.80%	96.94%	95.68%	88.14%	28.51%	460x	3.94%
200105771-788-S14	98.32%	97.63%	96.81%	95.49%	77.60%	5.02%	353x	3.88%
200105772-789-S15	98.36%	97.75%	96.79%	95.38%	84.77%	17.21%	409x	3.92%
200105773-790-S16	98.00%	97.14%	96.00%	93.23%	89.15%	0.42%	277x	4.18%
200105774-791-S17	98.09%	97.44%	96.38%	94.34%	72.12%	1.76%	328x	3.99%
200105775-792-S18	98.21%	97.47%	96.48%	94.46%	76.65%	2.21%	341x	4.00%
200105776-793-S19	98.10%	97.48%	96.36%	94.05%	67.65%	0.77%	307x	3.98%

Los criterios de filtrado utilizados para los archivos completos de variantes fueron: i) Frecuencia en GenomAD ≤ 0.05 ii) Variantes sinónimas, *inframe*, *missense*, *frameshift*, *nonsense*, *no start*, *no stop*, *site splicing*, *intronic*, UTR y CNVS (mayor de 50pb). iii) *ClinVar Pathogenicity* iv) Fracción de variante: ≥ 15 v) Profundidad: 30 Los criterios en la clasificación de variantes se rigen a las estandarizadas por la ACMG 2017 y pueden ser categorizadas como: Patogénicas (P), probables patogénicas (PP), variantes de significado incierto (VUS), probables benignas (PB) y benignas (B).

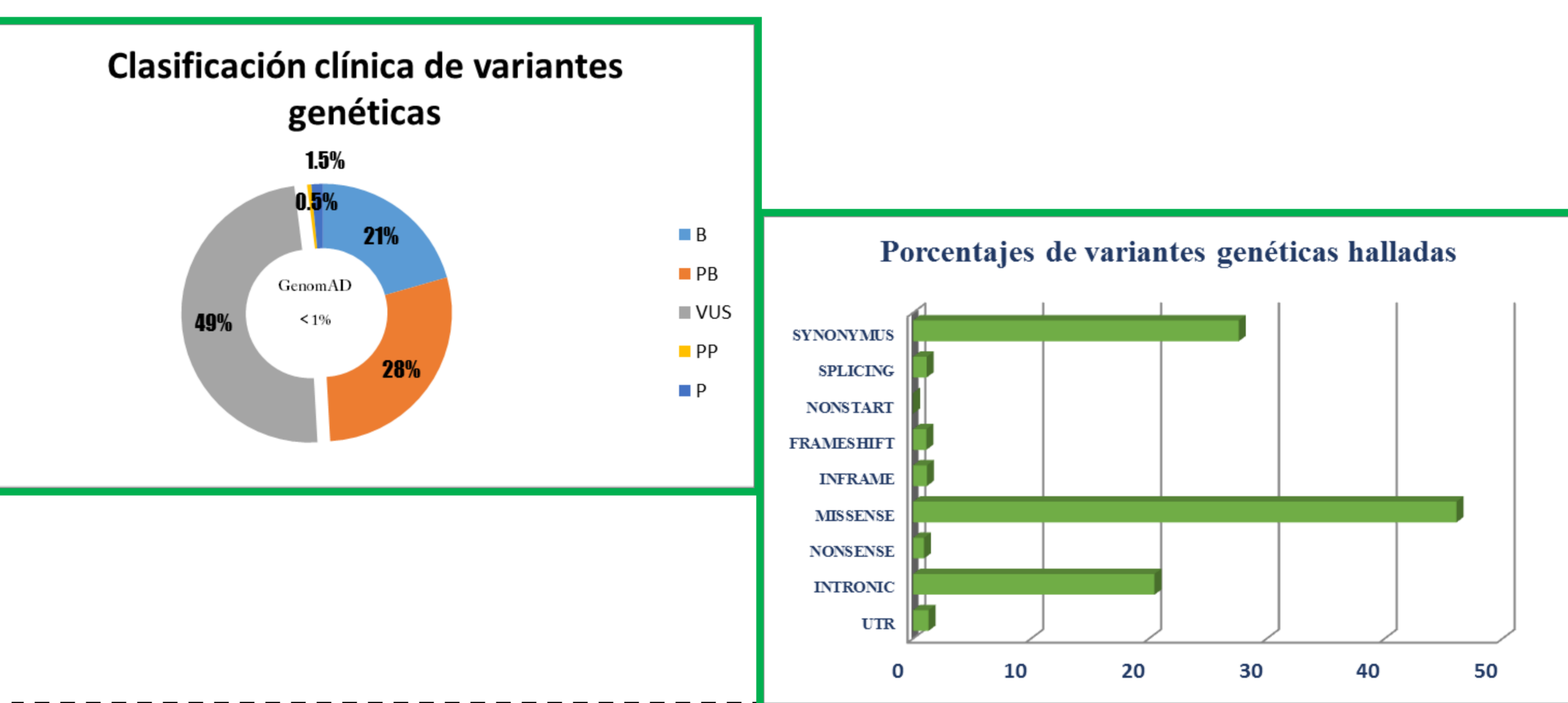


Resultados

Las variantes se filtran utilizando como primer criterio su frecuencia alélica en la población revisada en la base de datos de GenomAD. Las variantes clasificadas como benignas y probables benignas son las más abundantes así como las de significado incierto.

El porcentaje de variantes patogénicas en el universo de 870 casos evaluados a la fecha indica un 15% de pacientes con al menos una variante de este tipo relacionada con el fenotipo de la enfermedad genética.

Por otro lado, el tipo de variante más frecuente encontrada es la *missense* (cambio de aminoácido), siendo la variante *nonstart* la menos encontrada.



Conclusiones

- Se ha determinado el número y distribución de variantes genéticas en población peruana, considerando que INEN recibe pacientes de todo el territorio nacional.
- Consolidamos un modelo de flujo de trabajo Biólogo molecular-Médico Genetista que permite agilizar la resolución de casos y elaboración de informes médicos oportunos.
- El análisis de las variantes genéticas en un equipo multidisciplinario permite una clasificación precisa y enriquece la experiencia profesional, además expande el enfoque para la resolución de casos complejos.
- La contribución de los hallazgos en pacientes evaluados en el consultorio de genética permite establecer el diagnóstico definitivo de cáncer: cáncer hereditario, cáncer familiar y cáncer esporádico, haciendo posible el manejo de precisión del paciente de manera personalizada y planificando estrategias de prevención en la familias.

Referencias

1. Benson, D. A., Cavannah, M., Clark, K., Karisch-Hilzsch, I., Ostell, J., Pruitt, K. D., & Sayers, E. W. (2018). GenBank. *Nucleic Acids Research*, 46(D1), D41-D47.
2. den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., Roux, A. F., Smith, T., Antonarakis, S. E., & Taschner, P. E. (2016). HGVSC Recommendations for the Description of Sequence Variants: 2016 Update. *Human Mutation*, 37(6), 564-569.
3. Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipirala, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Muddipati, Z., Malheiro, A., McDaniell, K., Ovesky, M., Riley, G., Zhou, G., Holmes, J. B., Kattman, B., Maglott, D. R., ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018. Jan 4. PubMed PMID: 29165669.
4. Newman, V., Moore, B., Sparrow, H., & Perry, E. (2016). The Ensembl Genome Browser: Strategies for Accessing Eukaryotic Genome Data. *Methods in molecular biology* (Clifton, N.J.), 1375, 115-139.
5. Richards, S., Aziz, N., Bale, S., et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17, 405-423 (2015). <https://doi.org/10.1038/gim.2015.50>
6. Wang, Q., Pierce-Hoffman, E., Cummings, B. B., Altshuler, J., Franciosi, L. G., Gauthier, L. D., Hill, A. J., O'Donnell-Luria, A. H., Genome Aggregation Database Production Team, Genome Aggregation Database Consortium, Karczewski, K. J., & MacArthur, D. G. (2020). Landscape of multi-nucleotide variants in 125,748 human exomes and 15,708 genomes. *Nature communications*, 11(1), 2539.