DNAClinSUS v1: A Decision Support System to Identify Clinically Relevant Variants in Microarray Genotypes Using Panel-Based Reporting

Paulo Roberto Ferreira Júnior [1] , Lisane Brisolara de Brisolara [1] , Hugo Carlos Bolzon Gonzalez [1] , Sebastián Morales-Pison [2] , Ricardo Fernández-Ramires [2] , Vinicius Farias Campos [1] , and Flavio Fernando Demarco [1]

[1]Universidade Federal de Pelotas
Pelotas – Rio Grande do Sul – Brazil

[2]Universidad Mayor
Santiago – Chile

paulo.ferreira@inf.ufpel.edu.br

Abstract. *Translational genomics workflows still face a persistent operational gap between genotyping outputs and clinically usable, auditable interpretations: in many settings, results remain scattered across spreadsheets, ad hoc queries, and external websites, making reporting difficult to reproduce and to trace back to supporting evidence. DNAClinSUS v1 is a web-based clinical decision support system developed in the context of a SUS-oriented precision public health initiative in Brazil to address this gap by operationalizing a pragmatic genotype-to-report workflow in environments where patient genotypes are already available from microarray platforms. In its current version, DNAClinSUS v1 ingests single-individual CSV genotype files from microarrays, deterministically aligns SNP calls using array-manifest references, and supports panel-oriented interpretation through a locally managed knowledge base of genes and ClinVar-linked variants. The platform enables first-class curation and reuse of gene-based panels, supports SNP and polygenic panels via direct import of scoring files, and generates patient reports, as well as variant-level drill-down with links to external evidence resources. This paper presents DNAClinSUS v1 and discusses its constraints and requirements, data model, and the implemented workflow, emphasizing software engineering design choices and outlining its role as a concrete baseline that can be extended in future versions.*

1. Introduction

Precision medicine initiatives depend on integrating genomic evidence into routine clinical workflows. In practice, there is a persistent gap between raw genotyping outputs and clinically usable interpretations. This gap is especially relevant in public health settings, where scalable, auditable, and operationally simple tools are needed to support professionals in variant interpretation and diagnosis.

An additional challenge concerns the population specificity of genomic evidence. Much of the currently available evidence for variant interpretation and genetic risk estimation has been generated in predominantly European-ancestry populations, limiting its

direct transferability to highly admixed populations such as Brazilians. This is not only a matter of representation, but also of clinical validity, since allele frequencies, linkage patterns, local ancestry, and even the interpretation of variant pathogenicity may differ across populations, potentially affecting prioritization, classification, and downstream clinical decisions.

Recent work in Brazilian population genomics has reinforced this concern by highlighting both the underrepresentation of Brazil in global genomic resources and the presence of substantial Brazilian-specific genetic variation, including millions of variants not represented in major international databases [de Oliveira et al. 2024]. In parallel, national initiatives linked to the Brazilian Ministry of Health have framed large-scale genomic data generation as strategic infrastructure for the Brazilian Unified Health System (SUS) and for more precise, evidence-based public policies [Nunes et al. 2025].

The broader project in which this work is situated is motivated by population aging in Brazil and the need to build practical infrastructure to translate genomic evidence into public health and clinical decision-making workflows. In its conception, the initiative is anchored in a longitudinal research on aging (cohort) conducted in Brazil.

In this context, this aging cohort is not only epidemiologically relevant but also strategically important for genomic research in Brazil. The project foresees a future stage of genomic characterization of participants from that cohort and its expansion, enabling the analysis of genomic variants alongside detailed clinical, functional, and epidemiological data collected over time. This planned integration is expected to create a valuable opportunity to generate population-specific evidence on clinically relevant variants in a real-world Brazilian setting, helping reduce reliance on interpretation frameworks derived primarily from non-Brazilian populations.

DNAClinSUS v1 is presented in this paper as the first operational software layer of this broader precision public health initiative, which aims to integrate genomic and clinical information for use in the SUS. Within this strategic context, the contribution reported here addresses a foundational and immediately practical problem, namely, how to operationalize genotype-based clinical interpretation and panel-oriented reporting within a structured and traceable workflow. It was designed as an adaptable clinical decision support system that provides a simple and operational workflow focused on patient registration, panel curation, report generation, and direct variant traceability, while also providing a local infrastructure for storing genetic data from Brazilian individuals.

The problem addressed in this work is not transforming raw DNA sequencing reads into variants by aligning them to a reference genome, variant calling, and then conducting large-scale annotation. Instead, DNAClinSUS v1 is designed to consume the outputs of such bioinformatics pipelines as inputs for interpretation and reporting. In its current version, DNAClinSUS v1 supports a simpler genotype-to-report workflow based on microarray SNP (Single Nucleotide Polymorphism) genotypes [Gunderson et al. 2005]. The current version focuses on Illumina [Illumina, Inc. 2026] microarray-derived SNP genotype data. Importantly, this scope is a deliberate starting point rather than a hard limitation. The DNAClinSUS's extensible design is intended to serve as the foundation for future ingestion of sequencing-derived variant datasets (VCFs) as the broader project evolves. Thus, DNAClinSUS v1 addresses an immediate operational layer in our current

environment: transforming the deterministic SNP alignment of already available microarray genotypes, using ClinVar-supported interpretation metadata [Landrum et al. 2018], into clinical reports.

Positioning DNAClinSUS v1 relative to existing tools is also important. A relevant reference in the clinical variant interpretation space is VarSome Clinical, a commercial platform focused on curation and interpretation of genomic variants, supporting workflows such as exome/gene panel interpretation and multi-sample analyses [Saphetor SA 2026]. In practice, VarSome Clinical is typically used in a variant-centric manner, where users navigate and curate results by performing targeted searches and filters for variants (and, at most, by maintaining gene lists that reflect genes of interest). DNAClinSUS v1, in contrast, was designed around a panel-oriented workflow in which curated gene panels are first-class entities in the system and can be shared across users as part of a structured curation process; reports are then generated by applying these panels to patient data, rather than requiring each user to repeatedly re-specify variant queries. In addition, DNAClinSUS v1 supports the creation of SNP/polygenic panels [Lewis and Vassos 2020] via direct import of scoring files, enabling the operationalization of polygenic models at scale, including the ingestion of thousands of publicly available score definitions from international resources (e.g., polygenic score repositories) into a reusable, auditable panel structure.

Our tool was implemented as a web-based system that currently supports patient registration, genotype upload in CSV format (one individual per file), local management of genes and variants with ClinVar-supported interpretation [Landrum et al. 2018], creation and management of both gene-based panels and SNP/polygenic panels [Lewis and Vassos 2020], and generation of structured patient- and panel-specific reports with filtering and traceability to external genomic resources. These capabilities define the contribution of this paper: we present and discuss the design of an extensible clinical decision support system for genotype-based variant interpretation through panel-oriented reporting, and we demonstrate its feasibility as the first operational layer of a broader precision public health platform.

To present and discuss DNAClinSUS v1, the remainder of this paper is organized as follows. Section 2 describes the system constraints and requirements, motivating the primary design goals. Section 3 presents representative use scenarios supported by the current implementation and clarifies the practical scope of version 1. Section 4 details the underlying data model and the separation between reference knowledge, panel configuration, and patient/report execution data. Section 5 then describes the end-to-end system workflow and key user-facing components, illustrated by the main interface screens and report exploration features. Finally, Section 6 concludes the paper and outlines directions for future expansion.

2. Constraints and System's Requirements

DNAClinSUS v1 requirements were defined based on two general constraints: the system had to be usable in routine practice by a genetic consultant and extensible enough to serve as the software foundation of a broader precision public health platform.

A first design requirement was to make the workflow explicitly patient-centered. Instead of treating genotype files as isolated computational inputs, DNAClinSUS v1 or-

ganizes analysis around the patient as the primary entity. This decision supports repeated use over time, enables multiple panel-specific analyses for the same individual, and aligns the software structure with how clinical information is typically handled in practice. In engineering terms, this requirement also motivated separating patient data, panel configuration, and report outputs in the underlying data model, allowing the same patient to be associated with multiple analyses.

A second requirement concerned input standardization and operational simplicity. The system was designed to accept genotype files exported in CSV format from microarray platforms, with one file per individual and SNP genotype values represented as diploid calls (e.g., `AG`, `CT`, `TT`). By restricting the first version to a stable, narrow ingestion format, the implementation reduces parsing ambiguity, simplifies validation, and supports a more reliable genotype-to-report pipeline. Rather than attempting to support heterogeneous genomic formats from the outset, DNAClinSUS v1 prioritizes a platform that is already widely available and relevant in the target environment.

A third requirement was the adoption of a local genomic knowledge base for genes and variants, rather than relying solely on external live queries at report generation time. This requirement was motivated by performance, reproducibility, and operational stability. In practice, repeated remote queries may introduce latency, availability issues, or inconsistencies across analyses performed at different moments. By storing gene metadata and variant annotations locally—while preserving the linkage to external knowledge sources such as ClinVar—the system can support more predictable execution and more stable reporting behavior. This design also enables curation and reuse of previously imported entities, which is important in panel-oriented workflows with recurrent genes and variants.

Another system requirement was support for heterogeneous panel types within a single application model. DNAClinSUS v1 was therefore designed to support both gene-based panels (e.g., hereditary syndrome or disease-predisposition panels) and SNP/polygenic panels defined through imported scoring files (e.g., TSV-based specifications). From a software engineering standpoint, this requirement avoids fragmentation into separate tools for monogenic/hereditary and SNP-based workflows, while preserving a unified report-generation interface. It also anticipates future platform expansion by establishing a panel abstraction that can accommodate different interpretation strategies without requiring a complete redesign of the surrounding workflow.

Clinical auditability and traceability were also treated as system requirements. The system must expose the evidence used in interpretation, including ClinVar-derived metadata and classification labels [Landrum et al. 2018], and provide direct links to complementary resources such as ClinGen [Rehm et al. 2015], dbSNP [Sherry et al. 2001], and VarSome [Kopanos et al. 2019]. This requirement reflects the fact that clinical decision support tools must facilitate inspection and verification, not only automate output generation. In practical terms, this led to interface and data-organization decisions that preserve variant-level detail and support drill-down from report entries to external evidence sources.

Finally, extensibility was treated as a core architectural objective. Although DNAClinSUS v1 does not yet incorporate cohort data, it was designed to allow future integra-

tion of additional datasets and analytics services. This requirement is essential given the broader project context, in which the system is expected to evolve from a genotype-to-report module into a component of a larger precision public health infrastructure. Accordingly, the first version prioritizes clear separation of concerns across patient management, knowledge-base management, panel configuration, and report generation, creating a software baseline that can be expanded incrementally while preserving operational continuity.

3. Use Scenarios

To clarify the scope of DNACLinSUS v1, this section describes representative use scenarios that its implementation should support. The scenarios here emphasize the core properties of DNACLinSUS v1: a patient-centered organization, reusable panel definitions, a locally managed knowledge base of genes and ClinVar-linked variants, and auditable reporting with evidence drill-down.

A first use scenario is patient-specific genotype-to-report interpretation using a curated gene panel. In this scenario, the user registers a patient on the platform and uploads a single-individual genotype file generated by an Illumina microarray workflow. The user then selects a gene-based panel curated in the system (e.g., a hereditary-risk panel) and generates a report for that patient-panel pair. The report is presented as a structured table that can be filtered by gene and clinical significance, enabling a focused review of potentially relevant findings. This scenario captures the primary operational value of DNACLinSUS v1 in environments where genotyping results exist, but interpretation is otherwise fragmented across spreadsheets and multiple external resources.

A second use scenario is patient-specific reporting using SNP/polygenic panels imported from scoring files. In addition to gene-based panels, DNACLinSUS v1 supports the creation of SNP- and polygenic panels by importing scoring files. In this scenario, the user selects an imported SNP/polygenic panel and generates a report for a registered patient, producing panel-specific outputs based on the SNP genotype calls available in the uploaded microarray file. This scenario illustrates how the platform operationalizes a second interpretation paradigm within the same workflow, without requiring separate tools or separate data representations.

A third use scenario is shared panel curation and reuse across analyses. Panel-oriented workflows assume that panels are reusable organizational artifacts rather than ad hoc queries that each user repeats. DNACLinSUS v1 supports this by allowing gene-based panels and SNP/polygenic panels to be created once and then applied consistently across multiple patients. In practice, this enables a structured curation process in which panel definitions represent shared institutional knowledge and can be refined over time, improving consistency and reducing repeated manual specification effort during routine report generation.

A fourth use scenario is audit-oriented review with evidence drill-down for clinical verification. After a report is generated, users can inspect individual report entries and open the variant detail view, which exposes ClinVar-derived metadata and provides direct links to the following complementary external resources: ClinGen, dbSNP, and VarSome. This scenario highlights the non-black-box behavior of DNACLinSUS v1: the system not only surfaces results, but also preserves traceability to the evidence sources that support interpretation, enabling second-level review and verification in decision-support contexts.

Taken together, these scenarios show that DNACLinSUS v1 already supports a complete and operational genotype-to-report workflow for microarray-derived SNP genotypes, including both gene-based and SNP/polygenic panels, shared panel curation, and auditable report exploration with evidence drill-down. The current implementation should therefore be understood as a deployable workflow layer for interpretation and reporting, designed to support immediate practical use and to serve as a foundation for future extensions in the broader SUS-oriented precision public health initiative.

4. Data Model

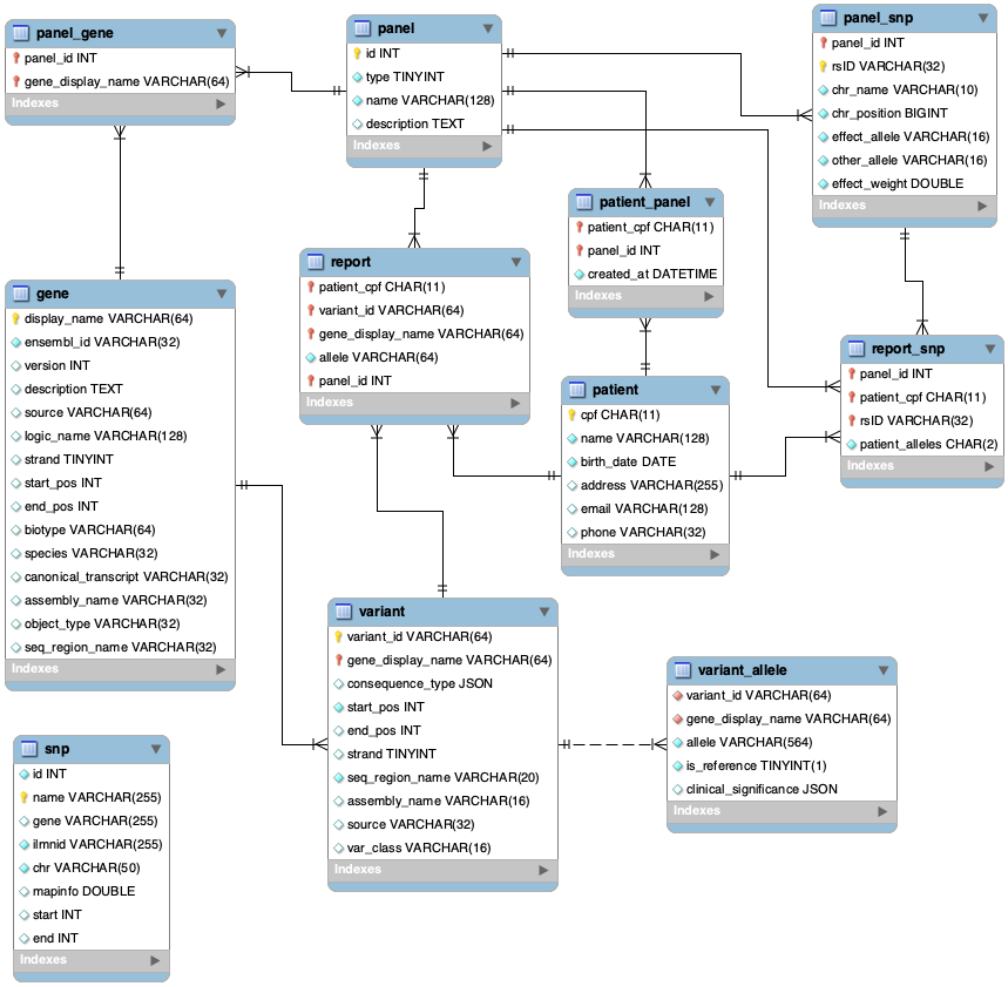


Figure 1. Relational Model of the DNAClinSUS v1 database schema.

The data model of DNAClinSUS v1 was designed to support a patient-centered, panel-oriented genotype interpretation workflow while preserving a clear separation between curated genomic knowledge, panel configuration, and patient-specific data. This separation is central to the maintainability and extensibility of the system because it allows the same reference knowledge and panel definitions to be reused across multiple patients and reports. Figure 1 presents the Relational Model diagram of the implemented database schema.

At a high level, the schema can be interpreted as comprising three major components: (i) a reference knowledge component, responsible for storing locally curated genomic entities, array reference SNP definitions, and interpretation metadata; (ii) a panel configuration component, responsible for defining reusable analysis structures; and (iii) a patient-reporting component, responsible for storing patient data and report outputs generated for clinical use.

The reference knowledge component implements the local genomic knowledge base used during report generation and genotype alignment. In the current schema, this component is centered on tables such as `gene`, `variant`, `variant_allele`, and `snp`. The `gene` table stores gene-level metadata (e.g., identifiers, genomic region, assembly, and related attributes), while `variant` stores variant-level records associated with genes, including source-related metadata imported from external resources (primarily ClinVar in the current implementation). The `variant_allele` table stores allele-specific interpretation information, including reference-allele flags and clinical significance labels. Together, these tables separate reusable genomic reference knowledge from patient-specific genotype observations, enabling consistent interpretation and reuse across many patient analyses.

In addition, we have created an auxiliary `snp` table that stores reference SNP definitions derived from the Illumina Global Screening Array (GSA) manifest files (CSV/BPM), which are publicly distributed as support files by Illumina [Illumina 2024], representing the set of loci expected in the microarray platform. In the microarray used as reference in DNAClinSUS v1, this corresponds to approximately 650,000 loci. This table is operationally important because it aligns patient genotype files with the SNP identifiers used by the system during ingestion and reporting. The genotype data is stored as a pure text file containing only the alleles.

The panel configuration component defines how curated knowledge is operationalized into reusable interpretation units. Its central table is `panel`, which stores panel metadata, including name, description, and context/type. Two complementary realization mechanisms are supported in the current version. For gene-oriented workflows, `panel_gene` represents the association between a panel and one or more genes of interest. For SNP/polygenic workflows, `panel_snp` stores SNP-level definitions and scoring-related attributes used in imported score-based models. This design provides a unified panel abstraction while preserving the flexibility required to support both hereditary/gene-panel and SNP/polygenic use cases in the same platform.

The patient-reporting component captures the system's execution state during real use. The `patient` entity stores patient registration data and acts as the anchor for downstream analyses. The relation between a patient and a selected panel is represented by `patient_panel`, which supports repeated analyses and the generation of multiple reports for the same patient across different panel definitions. The `report` entity stores report records and interpreted findings associated with a given patient-panel context, while `report_snp` stores genotype values mapped to panel SNPs in SNP/polygenic workflows, based on alignment between uploaded genotype data and the reference SNP definitions maintained in `snp`.

5. System Workflow

At the implementation level, the use scenarios from Section 3 were mapped directly onto six functional modules described in this section. Figure 2 summarizes how each use scenario compounds each DNAClinSUS v1 implemented module.

Use scenario	Patient registry	Genotype ingestion	Genes/ClinVar managment	Gene panels	SNP/polygenic panels	Evidence drill-down
1. Gene panel			✓	✓		
2. SNP/polygenic panel					✓	
3. Patient reports	✓	✓		✓	✓	
4. Audit-oriented review						✓

Figure 2. Mapping representative use scenarios to DNAClinSUS v1 modules.

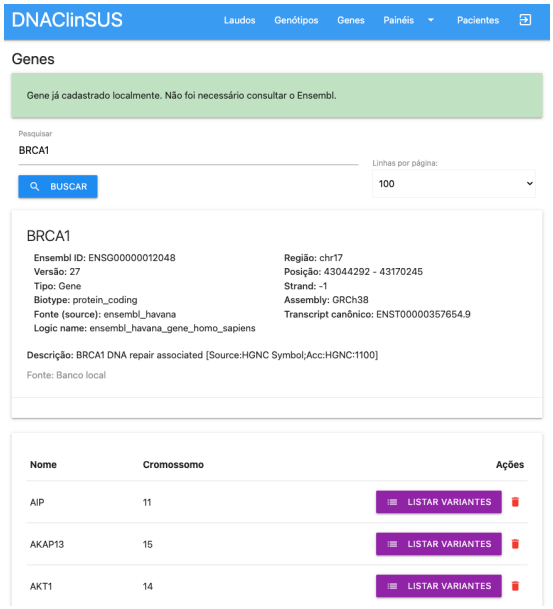
The DNAClinSUS v1 workflow starts with patient registration, followed by genotype upload (CSV, one individual per file, SNP-based Illumina microarray output). Then it proceeds to panel-oriented interpretation using a locally managed knowledge base of genes and variants. The final output is a structured report associated with a specific patient and panel, with support for filtering, inspection, and variant-level drill-down to external evidence sources. This organization reflects our concern to handle the constraints and requirements discussed in Section 2, providing an integrated operational path from input data to clinically navigable results.

The patient management module acts as the anchor of the workflow. Patients are registered on the platform and serve as the central entity for downstream analyses, enabling multiple reports to be generated for the same individual across different panels. A dedicated genotype ingestion module allows the user to upload the patient's microarray genotype file (CSV, one individual per file) and associate it with the corresponding patient record, ensuring that subsequent panel-based analyses and reports are generated from the correct genotype dataset. This software design supports how genomic interpretation is handled in practice, where the same patient may be evaluated under different clinical hypotheses and panel definitions over time.

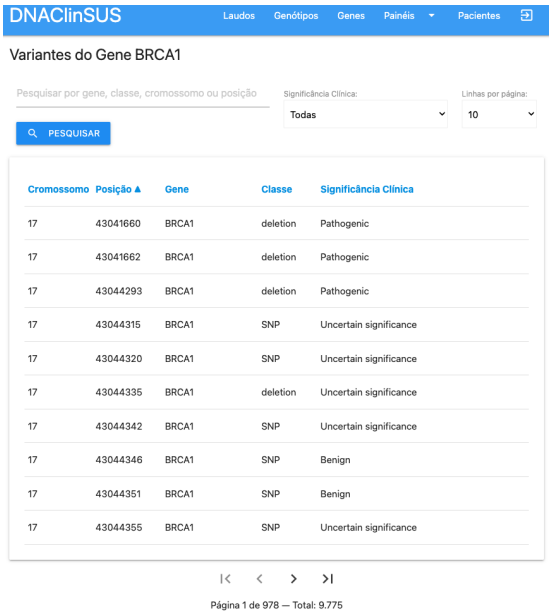
The gene and variant knowledge base module provides local management of genomic entities used during report generation. The interface supports gene search and import, and indicates when a gene is already stored locally, avoiding unnecessary repeated external queries. This behavior reflects the use of a local persistence layer for gene metadata (e.g., Ensembl identifiers, genomic coordinates, assembly, and transcript-level information). The same module also supports loading and listing gene variants, with ClinVar as the primary source of variant interpretation metadata. These locally managed entities are subsequently reused across panels and reports, improving consistency and execution stability. Representative interfaces for gene lookup and variant loading are shown in Figure 3a and Figure 3b.

Panel management is one of the central architectural features of DNAClinSUS v1 because it supports two distinct but complementary interpretation paradigms through two distinct modules. The first is gene-based panel definition, used for hereditary syndrome and disease-predisposition workflows, in which one or more genes of clinical interest are

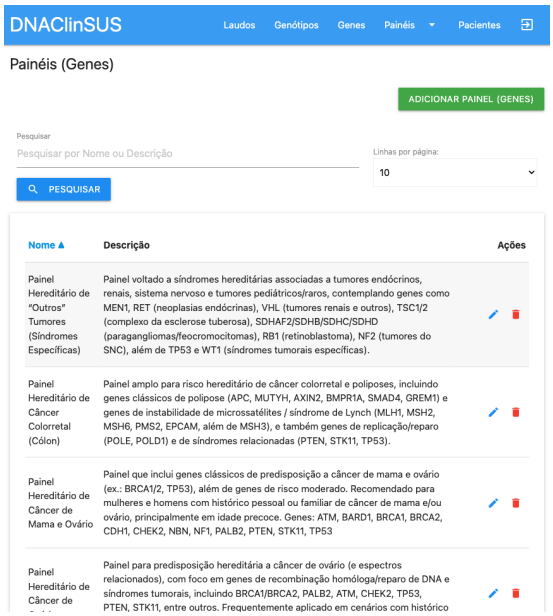
grouped into a curated panel. The second is the SNP/polygenic panel definition, created by importing scoring files to enable the representation of larger SNP-based models. Supporting both paradigms in the same platform avoids fragmentation into separate tools and preserves a consistent report-generation workflow independent of panel type. Figures 3c and 3d illustrate the panel management interfaces for both gene-based and SNP/polygenic configurations, respectively.



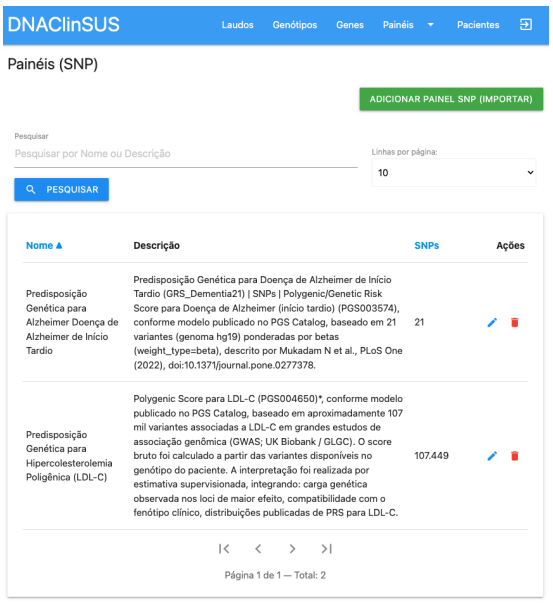
(a) Gene search and import interface with local persistence behavior.



(b) Variant loading/listing associated with genes.



(c) Gene-based panel management interface.



(d) SNP/polygenic panel management interface with scoring-file import.

Figure 3. Knowledge base and panel management interfaces in DNAClinSUS v1.

Report generation and exploration constitute the system's main interface. For a

given patient, DNAClinSUS v1 lists available panels and allows the user to generate a new report or open an existing one. Figure 4 shows the patient-panel report selection screen.

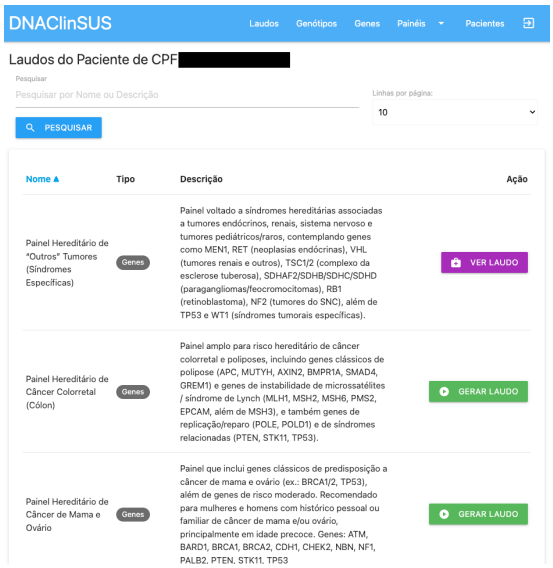
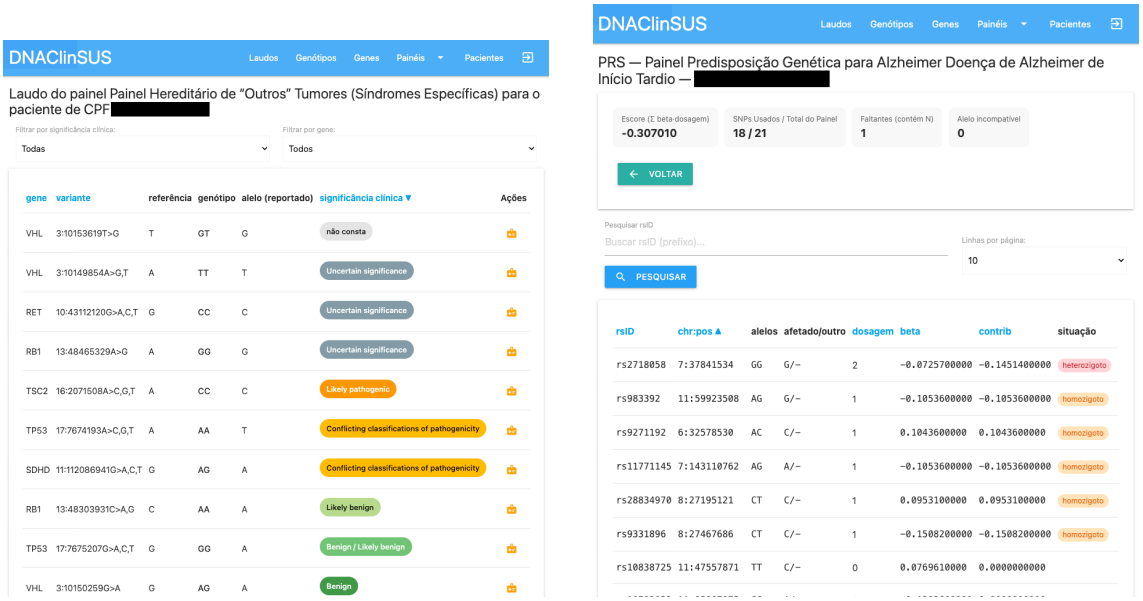


Figure 4. Patient-panel report selection interface (generate/view report).

Generated gene-panels based reports are displayed as structured tables containing, at minimum, the gene, variant, reference allele, patient genotype, reported allele, clinical significance, and action controls. The interface also provides filtering by gene and clinical significance, improving usability in review workflows by allowing users to focus on subsets of interest. Clinical significance labels follow ClinVar-style categories, including *Pathogenic*, *Likely pathogenic*, *Uncertain significance*, and *Conflicting classifications of pathogenicity*, as well as cases without a mapped classification. Figure 5a shows an example of the structured report view for a gene panel.

A key feature of the exploration workflow is the variant detail drill-down. From report entries, users can open a detailed variant view that presents ClinVar metadata, including accession, variant type, review status, genomic location, and last evaluation date [Landrum et al. 2018], along with direct links to ClinGen, dbSNP, and VarSome. This drill-down mechanism is important for clinical auditability because it preserves traceability between report-level results and the underlying evidence sources, while avoiding a black-box user experience.

In addition to gene-panel reports, DNAClinSUS v1 provides a dedicated report view for SNP/polygenic panels, as illustrated in Figure 5b. This interface presents an aggregated score computed as the sum of per-locus contributions (effect size $\times$ dosage), together with operational quality indicators that contextualize the result, such as the number of SNPs effectively used versus the total defined in the panel, the number of missing loci, and counts of incompatible allele encodings when applicable. Below the summary, the report provides a structured per-variant breakdown, enabling inspection of how each locus influenced the final score and supporting an audit-oriented review of polygenic score calculations.



(a) Gene-panel report view with filtering and ClinVar-style clinical significance labels.

(b) SNP/polygenic report view summarizing the score and per-locus contributions.

Figure 5. Report views in DNAClinSUS v1.

6. Conclusion

This paper presented DNAClinSUS v1, a web-based clinical decision support system that operationalizes genotype-based variant interpretation and panel-oriented reporting for microarray-derived SNP genotypes. The system addresses a practical gap in translational genomics workflows: turning existing genotype outputs into structured reports that preserve traceability to interpretation evidence, are organized around patients, and can be executed repeatedly across curated panels within a single operational environment.

In the current implementation, our system supports SNP alignment based on Illumina array-manifest references, local management of genes and ClinVar-linked variants, import of SNP/polygenic scoring files, curation and reuse of gene-based and SNP/polygenic panels, and report generation with filtering and variant-level drill-down linked to external evidence resources.

By structuring patient-centered data management, shared panel curation, and evidence traceability within a single environment, DNAClinSUS v1 provides a concrete baseline for the broader SUS-oriented precision public health platform under development. From a computing perspective, the paper offers a reusable engineering blueprint: its requirements, modular workflow, and data model can serve as a reference for other groups aiming to build computational tools around genomic data.

Future work in the broader project includes integrating additional clinical and epidemiological datasets, expanding data ingestion to support other genomics input formats, adding additional interpretation modules, and conducting formal evaluations of usability, performance, and decision-support impact in applied clinical trials.

Acknowledgments

The authors acknowledge financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (Chamada CNPq 33/2024 - Linha 2, Processo #406934/2024-0).

References

de Oliveira, T. C., Colli, L. M., Hünemeier, T., Pereira, L. V., et al. (2024). Population molecular genetics in brazil: From genomic databases and research to the implementation of precision medicine. *Human Genetics and Genomics Advances*, 5(3):100381.

Gunderson, K. L., Steemers, F. J., Lee, G., Mendoza, L. G., and Chee, M. S. (2005). A genome-wide scalable snp genotyping assay using microarray technology. *Nature Genetics*, 37(5):549–554.

Illumina (2024). Infinium global screening array v4.0 product files. `https://support.illumina.com/downloads/global-screening-array-v4-product-files.html`. Accessed: 2026-02-24.

Illumina, Inc. (2026). Illumina: Company information. `https://www.illumina.com/company.html`. Accessed: 2026-02-27.

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Aguilera, M. A., Meyer, R., and Massouras, A. (2019). Varsome: the human genomic variant search engine. *Bioinformatics*, 35(11):1978–1980.

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., Jang, W., Katz, K., Ovetsky, M., Riley, G., Sethi, A., Tully, R., Villamarin-Salomon, R., Rubinstein, W., and Maglott, D. R. (2018). Clinvar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, 46(D1):D1062–D1067.

Lewis, C. M. and Vassos, E. (2020). Polygenic risk scores: from research tools to clinical instruments. *Genome Medicine*, 12(1):44.

Nunes, K., Castro e Silva, M. A., Rodrigues, M. R., Lemes, R. B., Pezo-Valderrama, P., Kimura, L., de Sena, L. S., Krieger, J. E., Varela, M. C., de Azevedo, L. O., Camargo, L. M. A., Ferreira, R. G. M., Krieger, H., Bortolini, M. C., Mill, J. G., Sacuena, P., Guerreiro, J. F., de Souza, C. M. B., Veronese, F. V., Vianna, F. S. L., Comas, D., Pereira, A. C., Pereira, L. V., and Hünemeier, T. (2025). Admixture's impact on brazilian population evolution and health. *Science*, 388(6748):eadl3564.

Rehm, H. L., Berg, J. S., Brooks, L. D., Bustamante, C. D., Evans, J. P., Landrum, M. J., Ledbetter, D. H., Maglott, D. R., Martin, C. L., Nussbaum, R. L., Plon, S. E., Ramos, E. M., Sherry, S. T., Watson, M. S., and ClinGen (2015). Clingen—the clinical genome resource. *New England Journal of Medicine*, 372(23):2235–2242.

Saphetor SA (2026). Varsome clinical. https://landing.varsome.com/varsome-clinical. Accessed: 27 Feb. 2026.

Sherry, S. T., Ward, M.-H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., and Sirotkin, K. (2001). dbsnp: the NCBI database of genetic variation. *Nucleic Acids Research*, 29(1):308–311.