Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genomic landscape of inherited retinal diseases for gene therapy and clinical trials in Argentina

Marcela Ciccioli
Asociación Stargardt-APNES Retina Argentina
Jazmín Martínez
Administración Nacional de Laboratorios e Institutos de Salud (ANLIS) “Dr. Carlos G Malbrán”
Ezequiel Fernández
Fundación Actyon
Alejandra Antacle
Universidad Católica Argentina
Laura Echandi
Consultores Oftalmológicos
Daniel Hidalgo
Instituto Santa Lucía
Mario Saravia
Buenos Aires Macula Research
Rebeca Valero
Universitat de Barcelona
Miquel Tuson
Universitat de Barcelona
Gemma Marfany
Universitat de Barcelona
Roser Gonzàlez-Duarte
Universitat de Barcelona
Fabiana Arzuaga
Fundación Actyon
Hernán Dopazo
hernan.dopazo@ege.fcen.uba.ar

UBA-CONICET

Article

Keywords: Inherited retinal diseases, genetic diagnosis, next-generation sequencing, precision ophthalmology, clinical trials, advanced therapy medicinal products (ATMPs), Argentina

Posted Date: July 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9970614/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Inherited retinal diseases (IRDs) are highly heterogeneous disorders that pose challenges for molecular diagnosis and clinical translation. We analysed 808 molecularly solved patients from the SARA cohort, comprising 674 families and 1294 patient–variant observations involving 109 IRD-associated genes and 738 unique variants. Disease burden was concentrated in a limited number of diagnoses and recurrent genes, whereas extensive allelic diversity extended across a long tail of rare and structurally heterogeneous variants. This architecture was shaped by pleiotropy, diagnosis-specific genetic heterogeneity, and multiple modes of inheritance, including autosomal, X-linked, mitochondrial, and dual-mode patterns. Among unique variants, 529 had public clinical annotations, and 209 lacked public annotation. Reported variants accounted for most recurrent observations, whereas unannotated variants were enriched among partially resolved cases. These results define the genomic landscape of IRDs in Argentina and provide a framework for patient stratification in clinical trials and emerging gene-based therapies.

Introduction

Inherited retinal diseases (IRDs) constitute a clinically and genetically highly heterogeneous group of monogenic disorders characterised by progressive photoreceptor dysfunction or degeneration, ultimately leading to irreversible visual loss[1,2]. Given their frequently early onset, IRDs represent a major cause of visual disability in both pediatric and young adult ages, causing premature loss of occupational capacity and imposing a substantial socioeconomic burden on working-age individuals[3,4].

IRDs are broadly classified as non-syndromic or syndromic, depending on whether retinal degeneration occurs in isolation or in association with extra-ocular systemic manifestations[5]. To date, more than 500 IRD-associated genes have been identified, a number that continues to increase following a near-linear trend observed since 1989 for both syndromic and non-syndromic forms[6]. Modes of inheritance are diverse, with autosomal recessive (AR) transmission predominating (≈ 69% of loci, n = 323), followed by autosomal dominant (AD) inheritance (≈ 15%, n = 70). Additional loci show mixed AR/AD inheritance (≈ 7%, n = 36), while X-linked (≈ 4.7%, n = 22) and mitochondrial inheritance (≈ 3.6%, n = 17) account for smaller proportions[6].

Advances in genomic technologies have transformed IRD diagnosis, shifting from single-gene Sanger sequencing to targeted next-generation sequencing (NGS) panels and, increasingly, whole-exome (WES) and whole-genome sequencing (WGS) approaches[7]. Targeted gene panels remain particularly valuable, given the frequent involvement of non-canonical splicing sites in common IRDs[8].

In parallel, the therapeutic landscape for IRDs is rapidly evolving, driven by advanced molecular therapies that target underlying genetic mechanisms. These include DNA-based strategies such as gene replacement and genome editing, as well as RNA-based approaches involving splicing modulation, antisense oligonucleotides, and mRNA therapies[9–13]. The successful implementation of these strategies critically depends on accurate molecular diagnosis, underscoring the need for detailed clinical and genetic characterisation of affected individuals as a prerequisite for clinical trials and precision therapeutics.

Epidemiological and molecular studies of IRDs have been reported in numerous countries, including Italy, Spain, Portugal, Norway, England, Israel, China, Japan, Taiwan, Brazil, Mexico, and Chile[14–23]. Together, these studies highlight both the clinical relevance of IRDs and substantial regional differences in diagnostic coverage and access to molecular testing.

In Argentina, two major academic efforts have provided initial insights into the epidemiological and molecular landscape of IRDs. Mena et al.[24] characterised Stargardt disease (STGD) in 95 probands and 16 relatives, while Schlottmann et al.[25] reported the largest series to date, analysing 773 patients from 637 families across 13 provinces. Although these studies substantially advanced gene and variant identification, they also revealed persistent fragmentation and limited access to molecular diagnosis at the national level.

Globally, inherited retinal diseases (IRDs) have an estimated incidence of approximately 1:2000 and constitute one of the leading causes of severe vision loss among individuals aged 15–45 years[1]. Consistent with this, population-based data indicate that visual disability in Argentina constitutes a significant public health burden rather than a marginal condition. Nationally representative survey data indicate that visual impairment affects a substantial proportion of individuals aged ≥ 6 years in Argentina. Among the population reporting any disability (10.2% of individuals aged ≥ 6 years), approximately 25% present visual impairment, the vast majority of whom exhibit severe functional limitation; notably, 96.4% report great difficulty seeing or complete vision loss[26]. In this context, inherited retinal and optic nerve disorders—while individually rare—collectively contribute meaningfully to visual disability and warrant consideration within health-system planning frameworks.

Despite this burden, access to advanced diagnostic and therapeutic resources remains highly constrained in Argentina, reflecting limitations typical of many middle-income countries. Access to molecular genetic diagnosis is particularly restricted, owing to high costs, limited reimbursement, and a shortage of clinicians with expertise in retinal genetics. As a result, many patients experience prolonged diagnostic odysseys or remain without definitive molecular diagnoses, effectively excluding them from clinical trials and emerging gene-based therapies.

This gap underscores the need for coordinated, patient-centred models integrating clinical and molecular registries, public and regulatory oversight, nonprofit translational actors, and international collaboration to promote equity and reduce avoidable disparities.

Within this context, the SARA cohort emerged from a long-standing collaboration between the SARA Association (Stargardt APNES–Retina Argentina), a national patient advocacy organisation for individuals and families affected by inherited retinal diseases, and a network of academic, clinical, and translational research institutions. Rather than constituting an isolated registry, the SARA cohort was developed within a deliberately constructed institutional framework designed to overcome structural barriers to genetic diagnosis and research in Argentina.

The Actyon Foundation played a central role in conceiving, funding, and governing this framework, acting as a non-profit translational institution responsible for database architecture, data governance, and the establishment of sustained scientific collaborations. Through formal collaboration with the SARA Association, which facilitated patient recruitment, longitudinal engagement, advocacy activities, and clinical follow-up, and through institutional agreements with Argentina's national scientific system—including CONICET investigators, universities, and public research institutes—this framework enabled the systematic integration of clinical data, population-aware genomic analysis, and scientific production.

This institutional architecture transformed a patient-driven registry into a nationally curated research resource, supporting not only molecular diagnosis and clinical management but also the generation of reproducible scientific evidence. Within this translational setting, patients gained access to molecular diagnosis, research participation, and interventional trials, as exemplified by the LHON gene-therapy experience summarised in Box 1.

In this study, we present a comprehensive analysis of the genetic variant spectrum and phenotypic diversity observed in patients assessed by the SARA Association since 2013. The SARA Database, officially registered with the Ministry of Justice of the Argentine Republic, constitutes a systematically curated repository of clinical and molecular data encompassing diverse phenotypes, variant classes, zygosity states, and inheritance patterns. Collectively, these features position the SARA cohort as a national reference resource for characterising the genetic architecture of IRDs in Argentina and informing precision diagnostic and therapeutic strategies.

Box 1. Enabling gene therapy through the SARA framework

Inherited retinal and optic neuropathies are among the first monogenic disorders for which gene-based therapies have reached clinical application. Within the SARA framework—conceived as a national, patient-centred clinical and molecular resource—genetically resolved patients were systematically identified, stratified, and evaluated for eligibility for advanced therapeutic interventions.

A first milestone for this model was achieved in Leber hereditary optic neuropathy (LHON), most commonly caused by pathogenic variants in the mitochondrial MT-ND4 gene, particularly the m.11778G > A substitution. Between 2017 and 2018, ten patients (aged 9–33 years) with molecularly confirmed LHON were identified through the SARA registry and evaluated for a clinical trial: A Single Intravitreal Injection of rAAV2-ND4 for the Treatment of Leber's Hereditary Optic Neuropathy (ClinicalTrials.gov identifier: NCT03153293), sponsored by Huazhong University of Science and Technology. Patient stratification was enabled by the structured registry, expert clinical review, and a systematic assessment of the scientific and regulatory framework governing the trial[27]. Ethical and regulatory oversight included evaluation by the Advisory Commission on Regenerative Medicine and Cellular Therapies of the Argentine Ministry of Science (MINCYT).

Building on this experience, the SARA framework subsequently enabled access to approved retinal gene therapy. Following approval of voretigene neparvovec (Luxturna) by the Argentine regulatory authority (ANMAT) in 2021, patients with biallelic *RPE65*-associated retinal dystrophy were identified and evaluated within the cohort[28]. Since 2022, seven patients have been treated with Luxturna in Argentina, marking the first locally delivered advanced retinal gene therapy in the country and demonstrating the feasibility of translating population-aware genomic diagnosis into approved therapeutic action within a middle-income setting.

Beyond approved therapies, the SARA cohort currently supports evaluation and participation in multiple ongoing international phase III and II/III clinical trials, encompassing gene-specific approaches for Stargardt disease, gene-specific and gene-agnostic approaches for retinitis pigmentosa (*RPGR*, *RHO*-associated and agnostic forms), and mutation-specific *CEP290*-related disease, with additional trials under active consideration.

Together, these experiences illustrate how a coordinated framework—integrating patient organisations, non-profit translational foundations, clinical expertise, academic research, and public regulatory oversight—can transform genomic diagnosis into sustained therapeutic engagement. Rather than isolated successes, the SARA experience demonstrates a scalable, population-aware model for enabling equitable access to gene-based therapies as precision ophthalmology continues to expand.

Results

A total of 871 individuals with suspected inherited retinal diseases (IRDs) underwent genetic testing through the SARA programme. Among them, 808 patients had a positive molecular diagnosis, 29 had a negative result, and 34 had an inconclusive or unresolved result. These individuals originated from 674 normalised family units, including 577 singleton cases and 97 reported families, one of which had 7 affected individuals. The 808 molecularly solved patients (92.8% of all tested individuals) constitute the analytical cohort for all subsequent analyses presented in this work, encompassing clinical diagnoses, age of onset, zygosity categories, IRD gene frequencies, genotypes, and variant characteristics. The application of the rule-based framework classified 635 patients (78.6%) as resolved and 173 patients (21.4%) as partially resolved.

Clinical diagnoses. Figure 1, panel A, shows the distribution of the 12 diagnoses with a prevalence of at least 1% in the cohort. Retinitis pigmentosa (RP) predominated, accounting for 40.6% (328 patients), followed by Stargardt disease (STGD) at 19.7% (159 patients). Together, these two conditions represented nearly 60% of the cohort.

Other major inherited retinal diseases were less frequent, including Leber congenital amaurosis (LCA; 7.4%, 60 cases) and Leber hereditary optic neuropathy (LHON; 6.9%, 56 cases). The remaining diagnoses each contributed less than 4% of cases, comprising a long tail of rarer conditions such as Usher syndrome (USH), choroideremia (CHM), cone–rod dystrophy (CD/CRD), dominant optic atrophy (DOA), achromatopsia (ACHM), Bestrophinopathies (BEST), macular dystrophy (MD), and congenital stationary night blindness (CSNB).

Diagnostic categories with frequencies below 1% comprised 16 phenotypes affecting 47 patients (5.8%), including X-linked retinoschisis (XLRS), Bardet–Biedl syndrome (BBS), isolated optic atrophy (OA), combined CD/CRD/LCA presentations, albinism (ALB), Knobloch syndrome, recessive optic atrophy (ROA), Alström syndrome (ALMS), aniridia (ANIR), blue cone monochromacy (BMC), macular dystrophy North Carolina, Leigh syndrome, Stickler syndrome, familial exudative vitreoretinopathy (FEVR), combined optic atrophy/Stargardt disease (OA/STGD), and Refsum disease. Collectively, these diagnoses highlight the extensive phenotypic diversity captured within the SARA cohort (see Supplementary Table S1).

To further characterise the cohort of genetically resolved individuals, demographic and clinical parameters were analysed. Age at clinical onset ranged from early infancy, recorded as 0 years (1 month), to late adulthood (73 years). The median age of onset was 19 years, with a mean of 25.0 years. The distribution was bimodal, with the most frequent ages of onset at 15 and 17 years, and interquartile values spanning from 11 years (Q1) to 38 years (Q3). Overall, 415 patients (51.4%) manifested symptoms before 20 years of age, and 429 patients (53.1%) by 20 years of age, indicating a marked concentration of cases during childhood and adolescence. The full distribution is presented in Supplementary Fig. 1.

Additionally, we examined age of onset across diagnostic categories. The earliest recorded ages of onset were observed in LCA and RP, both of which included patients with onset in infancy (0 years), although their overall distributions differed substantially. LCA showed a younger age profile, with a median onset of 11.5 years, whereas RP showed a broader, later distribution, with a median onset of 32 years. Stargardt disease (STGD) showed onset from early childhood (minimum 1 year; median 16 years), while cone/cone–rod dystrophy (CD/CRD) and macular dystrophy (MD) tended to appear later, with median ages of 22 and 35 years, respectively. These diagnosis-specific patterns are shown in Supplementary Fig. 2.

Gene-level architecture. Figure 1, panel B, summarises the distribution of disease-causing genes among the 808 genetically solved patients, calculated at the patient–gene level and therefore not expected to match one-to-one with phenotype-level diagnostic counts. The figure highlights the 23 genes that individually accounted for at least 1% of diagnoses. *ABCA4* was by far the most frequent contributor, involved in 20.3% of molecular diagnoses (164 patients). This count includes the majority of Stargardt disease (STGD) cases, as well as patients with other clinical diagnoses, such as retinitis pigmentosa (RP), illustrating that gene-level and phenotype-level summaries capture complementary dimensions of the cohort. The following most common genes were *RPGR* (8.4%, 68 patients), *USH2A* (6.6%, 53 patients), and the mitochondrial gene *ND4* (5.7%, 46 patients). Other recurrent contributors included *CHM* (3.1%, 25), *CRB1* (2.6%, 21), *OPA1* (2.4%, 19), *RPE65* (2.2%, 18), *BEST1*, *PRPH2*, and *RHO* (2.1%, 17 each), *EYS* (1.9%, 15), *CEP290*, *NR2E3*, and *RDH12* (1.7%, 14 each), *PROM1*, *PRPF31*, and *RP2* (1.6%, 13 each), *RP1L1* (1.4%, 11), *PRPF8* and *RP1* (1.2%, 10 each), and *CLN3* and *GUCY2D* (1.1%, 9 each).

Figure_1.png here.

Figure 1. Clinical and gene-level architecture of the SARA cohort. (A) Distribution of inherited retinal disease (IRD) diagnoses with a prevalence at or above 1% in the cohort. Bars indicate the number of patients, with percentages calculated among genetically solved cases. (B) Recurrently implicated genes with frequencies at or above 1% among genetically solved patients, calculated at the patient–gene level. Patients carrying variants in multiple IRD-associated genes could contribute to more than one gene category. Genes with frequencies below 1% are reported in Supplementary Table S2.

Taken together, these recurrent genes accounted for most genetically solved cases, although the overall gene landscape remained highly heterogeneous, with 86 genes detected in fewer than 1% of patients. These low-frequency contributors are displayed separately in

Supplementary Table S2, highlighting the extensive genetic diversity characteristic of IRDs. In addition, 21 patients carried variants in more than one IRD-associated gene, including secondary or non-primary gene findings (Supplementary Table S3).

Genotype–inheritance patterns. Figure 2 summarises 831 patient–gene configurations derived from the 808 genetically solved patients. Each configuration represents the inheritance-compatible genotype pattern assigned to a given gene in a given patient, based on zygosity, number of variants per gene, patient sex, and the established mode of inheritance for each gene. The operational pattern definitions and gene-level inheritance assignments are provided in Supplementary Tables S4 and S5. The difference between the number of patients and the number of configurations reflects individuals with findings in more than one IRD-associated gene, including secondary or non-primary gene findings that were retained in the dataset but did not necessarily explain the patient's primary clinical diagnosis.

This framework captures how pathogenic or candidate disease-associated variants combine within individuals in accordance with expected Mendelian patterns, alongside recognised non-Mendelian mechanisms such as mitochondrial inheritance. Autosomal recessive (AR) inheritance was the predominant category, accounting for 435 configurations (52.3%), and was mainly driven by biallelic compound heterozygous genotypes (371 configurations), followed by homozygous AR genotypes (60 configurations). The four AR heterozygous carrier configurations corresponded to single-variant findings in AR genes and were retained as secondary or partial genetic evidence.

Autosomal dominant (AD) configurations represented the second most frequent inheritance mechanism, comprising 182 configurations (21.9%). X-linked (XL) configurations accounted for 120(14.4%), mostly corresponding to hemizygous males (105). The remaining 15 configurations occurred in heterozygous females carrying variants in X-linked IRD genes. These heterozygous female configurations were retained as a distinct category because they may be clinically relevant in affected females, potentially through skewed X-chromosome inactivation or other mechanisms that affect mutant allele expression. Mitochondrial (MT) findings formed a distinct category, accounting for 58 configurations (7.0%), consistent with likely homoplasmic pathogenic variants.

Finally, genes with established dual modes of inheritance (AR/AD genes) represented 36 configurations (4.3%). Within this group, 25 configurations involved single heterozygous variants in genes that can act through either dominant or recessive mechanisms depending on molecular and phenotypic context, whereas 11 configurations were compatible with biallelic involvement.

Overall, genotype–inheritance patterns were largely consistent with Mendelian expectations, while also encompassing recognised non-Mendelian mechanisms, most notably mitochondrial inheritance.

Figure_2.png here.

Figure 2. Genotype–inheritance patterns in the SARA cohort. Stacked bars represent 831 patient–gene configurations assigned to inheritance-compatible genotype patterns. Individual patients may contribute more than one configuration when variants were identified in multiple IRD-associated genes, including secondary or non-primary findings.

As an internal data quality control, we assessed the consistency between the modes of inheritance assigned in the SARA cohort and the curated classifications provided by the RetiGene Atlas[6]. For each patient-gene configuration, the gene-level inheritance inferred in SARA was systematically compared with the corresponding RetiGene annotation for inherited retinal diseases. *RBP4* showed a formal gene-level difference, being classified as AR/AD in SARA but AR in RetiGene; however, the two SARA configurations involving *RBP4* were biallelic and therefore compatible with recessive disease expression.

In addition, nine genes observed in the SARA cohort were not present in the current RetiGene reference used here: *COL11A2*, *CPLANE1*, *PITPNM3*, *PRDM13*, *RDH11*, *RIMS1*, *ROM1*, *RP9*, and *UNC119*. These genes were rare in the SARA cohort and together were observed in 12 patients.

Although these genes are not included in the current RetiGene reference used here, pathogenic or likely pathogenic variants in some of them were observed in our cohort. Further work on functional validation of the identified variants, as well as reports of other families from different genetic backgrounds carrying the same or other potentially pathogenic variants, may help clarify whether they should be considered causative, modifier, historical, or non-primary annotations in IRD.

Based on the RetiGene annotations, which assign each gene to syndromic, non-syndromic, or dual (both) disease categories, we examined the distribution of cases across these clinical classifications among the 100 genes shared between the SARA cohort and RetiGene Atlas. The majority of cases were associated with genes classified as non-syndromic (n = 564), while a substantial proportion mapped to genes annotated as having both syndromic and non-syndromic presentations (n = 209). Genes classified as exclusively syndromic accounted for fewer cases (n = 46). The predominance of non-syndromic cases, together with a sizeable contribution from dual-phenotype genes, is consistent with the known genetic architecture of inherited retinal diseases and reflects the pleiotropy of many IRD-associated genes.

Molecular variants. Figure 3 shows the molecular annotation landscape of the 738 unique variants identified in the SARA cohort. At the DNA level, the variant spectrum was dominated by single-nucleotide variants (SNVs; 548, 74.3%), followed by deletions (153, 20.7%) and duplications (30, 4.1%), with insertions (7, 0.9%) representing a smaller fraction. At the predicted protein or transcript-effect level, missense substitutions were the most frequent consequence (361 variants, 48.9%), followed by frameshift variants (134, 18.2%), stop-gain variants (95, 12.9%), and splice-altering variants (77, 10.4%).

The unknown protein-effect category (54 variants, 7.3%) mainly captured variants for which a precise amino-acid consequence could not be assigned from the available annotation. This group was enriched for exon-level deletions, duplications, and other structural events, as well as intronic or non-coding annotations whose impact is better represented at the DNA or transcript level than by a single protein change.

This distribution is consistent with the long-tail recurrence spectrum typically observed in inherited retinal disease cohorts, in which most variants are observed only once or a few times (often referred to as “private” mutations, as they are usually confined to a single family), whereas a small number of well-established alleles account for a disproportionate fraction of observations. These recurrent variants likely represent ancient mutations that have disseminated across populations and continents (Supplementary Fig. 3). Together, the DNA- and protein-level summaries provide the molecular context for the integrative analyses presented below.

Figure_3.png here.

Figure 3. Distribution of DNA variant types and predicted protein-level consequences across unique variants identified in the SARA cohort. SNVs and missense changes predominated, with a substantial contribution from loss-of-function variants, including frameshift and stop-gain alterations. Variants lacking an annotated protein-level consequence were classified as Unknown. Synonymous variants (n = 3) were retained in the source tables but omitted from the protein-level display.

Figure 4 provides an integrated view of the molecular and clinical structure of the SARA variant landscape by jointly considering DNA-level variant class, protein-level effect, and ClinVar pathogenicity category, analysed separately for reported (R) and novel (N) variants as defined by VEP/ClinVar public clinical annotation. Each cell represents a distinct variant class, with counts reflecting unique alleles rather than patient-level observations, allowing direct comparison of annotation patterns independent of recurrence frequency. Marginal totals above each DNA class and along the right-hand side of each panel summarise the displayed non-synonymous variant categories across DNA-level classes and protein-level effects.

By definition, variants classified as novel in this framework lacked public clinical annotation recovered by VEP/ClinVar. Among the 529 reported variants, 368 (69.6%) carried ClinVar pathogenic or likely pathogenic (P/LP) evidence, 129 (24.4%) were classified as VUS, 30 (5.7%) as benign or likely benign (B/LB), and 2 (0.4%) as other or conflicting. This separation highlights the large fraction of the SARA allelic spectrum that remains outside public clinical variant interpretation resources.

Reported variants exhibited a highly structured and canonical profile. The majority were single-nucleotide variants (SNVs) associated with well-characterised protein-level consequences, including missense, frameshift, and stop-gain changes. They were strongly enriched for ClinVar pathogenic and likely pathogenic classifications, consistent with extensive prior clinical curation. This pattern is consistent with the cumulative effect of historical discovery, clinical validation, and repeated observation across studies and populations.

In contrast, variants lacking public clinical annotation, as recovered by VEP/ClinVar, occupied a markedly different region of the annotation space. While SNVs remained frequent, there was a relatively higher contribution of non-SNV events and a substantial fraction of variants with unknown or non-canonical protein-level effects. Notably, rare and atypical combinations of DNA and protein effect classes—largely absent from the reported variant set—were concentrated in the novel category, highlighting regions of the IRD variant landscape that remain underrepresented in public clinical interpretation resources.

Figure_4.png here.

Figure 4. Heatmap representation of the molecular annotation landscape of reported and novel variants. Variants are stratified by DNA variant type, protein-level effect, and ClinVar classification. Each cell corresponds to a distinct variant category, with values indicating counts of unique variants rather than patient-level observations. Facet headers report full DNA-class totals; when synonymous variants were omitted from the protein-level display, the number of displayed non-synonymous variants is shown in parentheses. Right-hand marginal totals indicate the number of displayed variants within each protein-level effect category. Synonymous variants were retained in the source tables but omitted from this protein-level display. Darker shading denotes higher variant counts.

At the case level, this distinction was reflected in molecular-resolution status. Resolved cases were more likely to contain at least one reported variant than partially resolved cases (533/635, 83.9% vs 131/173, 75.7%; Fisher’s exact test, p = 0.0182), whereas partially resolved cases were more likely to contain at least one novel variant (72/173, 41.6% vs 180/635, 28.3%; Fisher’s exact test, p = 0.00116). Consistently, resolved cases carried a higher mean number of reported variants per patient (1.30 vs 1.10; Wilcoxon $p = 9.53 \times 10^{-4}$), while partially resolved cases

carried a higher mean number of novel variants (0.51 vs 0.30; Wilcoxon $p = 1.57 \times 10^{-4}$). These results indicate that public clinical annotation directly contributes to complete molecular resolution, whereas variants lacking public clinical annotation are disproportionately represented among cases that retain residual interpretative uncertainty.

At the gene level, these patterns resolve into a highly structured distribution in which recurrent reported variants are confined to a small number of canonical IRD genes, reflecting well-established disease associations (Supplementary Figure S4). Among recurrent reported variants observed in five or more patients, *ABCA4* contributed the largest number of alleles, including c.3386G > T (p.Arg1129Leu; 29 patients), c.5882G > A (p.Gly1961Glu), c.3113C > T (p.Ala1038Val), and c.5603A > T (p.Asn1868Ile; 12 patients). Notably, p.Gly1961Glu and p.Asn1868Ile have been described as hypomorphic *ABCA4* alleles associated with reduced penetrance and milder or later-onset retinal phenotypes, particularly when occurring in trans with more severe pathogenic variants[29]. Additional recurrent alleles included *RPGR* c.1345C > T (p.Arg449Ter) and *USH2A* c.2299delG (p.Glu767SerfsTer21), consistent with their known involvement in X-linked retinitis pigmentosa and Usher syndrome, respectively. In contrast, novel variants—although largely observed only once—were not uniformly distributed across genes, but instead tended to cluster within a limited subset of recurrent IRD-associated loci (Supplementary Figure S5).

The single most frequent variant in the cohort, *ND4* m.11778G > A (p.Arg340His), accounted for the highest proportion of affected individuals and families; however, this mitochondrial variant is associated with Leber hereditary optic neuropathy (LHON) and reflects targeted recruitment and analysis described separately in Box 1, rather than unbiased population recurrence. Together, these results indicate that variant-level recurrence in the SARA cohort is dominated by canonical pathogenic alleles in high-frequency disease genes, whereas novel variants contribute predominantly to the broader spectrum of genetic heterogeneity captured by the cohort.

To achieve a more comprehensive and homogeneous representation of inherited retinal diseases (IRDs) in Argentina, we integrated two complementary variant-level resources, SARA and the published dataset from Schlottmann et al. (2023), hereafter referred to as the SPG et al. dataset, into a unified analytical framework. Harmonization of clinical phenotypes, gene annotations, and variant-level attributes enabled direct comparison across cohorts while preserving source-specific information. This integrated approach revealed a highly concentrated structure of the Argentine IRD landscape, with the 20 most frequent clinical phenotypes accounting for 97.3% of patients with a defined phenotype, and the 20 most frequently implicated genes accounting for 71.6% of patients with at least one disease-associated gene (see: Supplementary Tables S6 and S7, respectively). Together, these results provide a robust and scalable reference for describing the genetic and clinical architecture of IRDs in Argentina and establish a foundation for downstream epidemiological, diagnostic, and therapeutic analyses in our country.

Discussion

The SARA cohort reveals a strongly skewed distribution across all major analytical dimensions. This skewness reflects intrinsic properties of inherited retinal diseases rather than artefacts of cohort composition. IRDs are characterised by extreme genetic diversity constrained within a limited set of biological pathways and clinical phenotypes, resulting in a pleiotropic and genetically heterogeneous architecture that becomes apparent when cohorts are analysed at sufficient scale. The present study provides such a scale within a population-aware framework from a middle-income country, based on 808 molecularly solved patients, 109 disease-associated genes, and 738 unique variants systematically re-annotated with VEP and interpreted alongside ClinVar-derived clinical classifications.

Most disease-associated genes with assigned clinical diagnoses were linked to a single phenotype (88/109 genes), whereas 21 genes (~ 19%) were associated with two or more distinct clinical diagnoses, indicating that pleiotropy is a recurring feature of IRDs rather than isolated exceptions. Conversely, clinical diagnoses varied substantially in their degree of genetic heterogeneity, ranging from conditions associated with only one or two genes to highly heterogeneous entities such as retinitis pigmentosa, which involved 52 distinct disease-associated genes. Extended analyses of the complete gene–phenotype architecture are provided in Supplementary Section 4.

At the phenotypic level, age at onset was predominantly early, with most individuals developing clinical symptoms during childhood or adolescence. Although 29 distinct clinical phenotypes were identified after excluding non-primary annotations, disease burden was highly concentrated: the five most frequent entities accounted for nearly 80% of cases. A comparable pattern was observed at the molecular level. While 109 disease-associated genes were implicated, just 23 recurrent genes were observed in nearly three-quarters of genetically resolved individuals, with *ABCA4* alone accounting for over one-fifth of the cases. In this context, it is worth noting that the carrier frequency of *ABCA4* pathogenic alleles is highest in European-ancestry populations (around 12%)[30].

This structured heterogeneity extended to inheritance patterns. Although autosomal recessive biallelic configurations predominated, all major Mendelian and non-Mendelian inheritance modes were represented, including autosomal dominant, X-linked, mitochondrial, and context-dependent AD/AR inheritance genes.

At the allelic level, most unique variants had public clinical annotations retrieved from VEP/ClinVar, and recurrent observations were dominated by canonical pathogenic alleles, whereas novel variants occupied a complementary space characterised by low recurrence, structural diversity,

and limited functional annotation. Their enrichment among partially resolved cases highlights a central interpretative gap: alleles lacking public clinical annotation are not merely additional variants, but a major source of residual uncertainty in molecular diagnosis. Together, these observations highlight how public clinical variant resources consolidate a core of recurrent, well-characterised alleles while leaving a substantial fraction of the IRD molecular landscape incompletely represented.

Integrating the SARA cohort with the SPG et al. dataset confirms a highly concentrated yet heterogeneous architecture of inherited retinal disease in Argentina. This organisation has direct implications for diagnostic strategy design, supporting targeted approaches optimised for prevalent genes while underscoring the need for flexible frameworks that accommodate rare, atypical, and unexpected findings.

Importantly, the high diagnostic yield and refined variant interpretation achieved in this study cannot be attributed solely to sequencing technology or gene content. Instead, they reflect the tight integration of genomic data with longitudinal clinical expertise. The SARA cohort illustrates how an iterative, clinician-led workflow—combining deep phenotyping, repeated clinical evaluation, and continuous reinterpretation of genetic findings—can resolve complex cases, reduce uncertainty around variants of uncertain significance, and accommodate the phenotypic overlap and age-dependent manifestations characteristic of inherited retinal diseases. Crucially, this approach enables nuanced interpretation of atypical or previously uncharacterised genotype–phenotype relationships that would likely remain unresolved in purely technology-driven pipelines based solely on previously reported, more prevalent pathogenic allele variants.

Beyond aggregate analyses, the breadth of patients, genes, and variant types captured in the SARA cohort provides a rich substrate for exploring genotype–phenotype relationships. Although a systematic investigation of such relationships was beyond the scope of the present study, selected cases illustrate the interpretative depth afforded by this dataset. One such case involves a patient with retinitis pigmentosa carrying a large homozygous deletion encompassing exons 22–29 of *USH2A*. While this deletion had previously been reported only in a heterozygous state, this represents, to our knowledge, the first homozygous case identified. Despite extensive loss of protein domains, the patient exhibited a predominantly retinal phenotype with only mild, non-progressive auditory involvement, underscoring that large structural variants may yield variable clinical outcomes depending on residual functional interactions.

A second illustrative case concerns a patient with autosomal recessive retinitis pigmentosa carrying a homozygous nonsense variant in *EYS* (p.Trp1817Ter), an allele previously reported only once in heterozygous state in South Asian populations. The variant was confirmed by independent resequencing at a depth exceeding 200×, excluding technical artefacts. Extended runs of homozygosity spanning approximately 0.8 Mb surrounding the variant supported homozygosity-by-descent as the most likely mechanism underlying homozygosity. This case highlights the importance of integrating variant-level findings with genome-wide context to accurately interpret rare pathogenic alleles in genetically diverse populations.

Beyond its analytical contributions, this study demonstrates the translational impact of a coordinated institutional framework operating in a resource-constrained setting. The SARA cohort emerged from sustained collaboration between a patient association, a non-profit translational foundation, and the academic scientific system, enabling equitable access to genetic diagnosis in a context where such testing is not routinely covered by health systems.

This translational dimension is exemplified by the advanced therapeutic initiatives coordinated within the SARA framework (Box 1), which illustrate how population-aware genomic interpretation—embedded within integrated clinical, institutional, and patient-centred governance structures—can directly enable access to gene- and mutation-specific therapies in middle-income settings.

More broadly, the SARA experience illustrates that addressing the extreme genetic and phenotypic heterogeneity of inherited retinal diseases requires not only comprehensive genomic analysis, but also durable clinical, institutional, and governance frameworks capable of translating molecular diversity into actionable precision medicine. As gene-based therapies continue to expand, such models will be essential to ensure equitable participation of underrepresented populations in the evolving landscape of precision ophthalmology.

Methods

Patients cohort and clinical data

The patient cohort and associated clinical data were collected from the database of the Stargardt APNES–Retina Argentina Association, which is registered with the Argentine Ministry of Justice and has been active since 2013. The database includes 1,460 registered patients, of whom 871 underwent genetic testing. Among these, 808 patients had a positive molecular diagnosis and constitute the genetically solved analytical cohort evaluated in this study. Comprehensive ophthalmic examinations, including fundus evaluation, best-corrected visual acuity (BCVA), optical coherence tomography (OCT), fundus autofluorescence, and full-field electroretinography (ffERG), were considered for clinical diagnosis when available. Clinical diagnoses were harmonised to a unified nomenclature for inherited retinal diseases while preserving source annotations. Records labelled as non-primary annotations were not used for primary diagnosis assignment when a clinically defined IRD diagnosis was available for the same patient.

Genetic analysis

Genetic analyses were performed between 2013 and 2026, predominantly using targeted next-generation sequencing (NGS) panel approaches focused on genes associated with inherited retinal diseases. Analyses and panel design were initially carried out at the DBGen Ocular Genomics Laboratory (Barcelona, Spain) and, from 2024 onwards, by Actyon Genomics SL, at the CNAG sequencing service (Centre Nacional d'Anàlisi Genòmica, Barcelona, Spain). Panel design and subsequent updates were developed collaboratively by DBGen, Stargardt APNES–Retina Argentina Association, Actyon Foundation and Roche Diagnostics. Earlier panel versions included 150 IRD-associated genes, whereas later versions expanded to 346 nuclear genes and selected intronic regions. The current version of the panel comprises 416 nuclear IRD-associated genes, selected intronic and regulatory regions (5' and 3' untranslated regions, UTRs) in 45 nuclear genes, and complete mitochondrial genome sequencing.

The majority of patients underwent genetic testing using these targeted panels, while a subset of cases entered the cohort with prior genetic testing performed independently, most commonly whole-exome sequencing, performed in external clinical laboratories. All available genetic results were integrated and interpreted within a unified analytical framework. For most patients, DNA was obtained from saliva using standard kits.

Bioinformatics and variant detection

Sequence data were processed using bioinformatic workflows optimised for inherited retinal diseases. Raw reads were aligned and processed according to the reference genome and pipeline version in use at the time of testing, and final publication-level coordinates were harmonised to GRCh38 when base-level mapping was available. Only properly paired and uniquely mapped reads were retained. Read groups were assigned, duplicate reads were marked using Picard tools, and local realignment around indels and base quality score recalibration were performed using Genome Analysis Toolkit (GATK) workflows applicable at the time of analysis.

Single-nucleotide variants (SNVs) and small insertions/deletions were called using GATK-based pipelines with quality filtering to retain high-confidence calls. Structural variants and copy-number changes were assessed using complementary read-depth and breakpoint-based approaches, including ExomeDepth and Manta when applicable. Mitochondrial variants were detected using a dedicated mitochondrial variant-calling workflow. Runs of homozygosity were analysed using PLINK to support interpretation of rare homozygous variants under autosomal recessive inheritance models.

Variant harmonisation and annotation

Because genetic testing was performed over an extended period and across evolving panel designs, all variant records were harmonised into a common analytical structure before downstream analysis. Patient identifiers, family identifiers, gene symbols, clinical phenotypes, zygosity, inheritance annotations, DNA-level changes, and protein-level consequences were standardised. Each unique variant was assigned a stable variant identifier, allowing repeated observations across patients and families to be collapsed into a unique-variant table while retaining all patient-level genotype observations.

To reduce heterogeneity introduced by historical annotations, variants with sufficient base-level coordinates and allele information were re-annotated using the Ensembl REST Variant Effect Predictor (VEP) /vep/human/region endpoint. Reannotation was performed on GRCh38 coordinates and included HGVS notation, canonical transcript information, predicted molecular consequence, variant class, population allele frequencies, and colocated public variant identifiers. Queryable variants were submitted in batches, and returned JSON objects were flattened into tabular fields for integration with the SARA genotype database. For each VEP result, transcript selection prioritised the expected disease gene and, when available, the previously curated transcript, followed by transcript-level impact and canonical annotation. VEP-derived fields were then integrated back into the genotype-level table.

Clinical significance was assessed using ClinVar-derived colocated variant annotations returned by VEP, including colocated identifiers, clinical-significance terms, and maximum reported allele frequency when available. These annotations were converted into harmonised pathogenicity flags, including pathogenic/likely pathogenic (P/LP), variant of uncertain significance (VUS), benign/likely benign (B/LB), and loss-of-function (LOF) categories. Variants with public clinical annotation recovered by VEP/ClinVar were classified as reported, whereas variants without such annotation were classified as novel for the purposes of this study. The final publication dataset comprised 1294 patient-variant observations and 738 unique variants, of which 698 were successfully mapped through homogeneous VEP reannotation, and 40 were retained as non-queryable structural or incompletely specified alleles requiring curated interpretation.

Variant pathogenicity evidence was harmonised into ACMG-compatible clinical categories, including P/LP, VUS, and B/LB, using ClinVar-derived annotations recovered by VEP together with predicted molecular consequence, allele frequency, and expert review where required for case-level interpretation. Loss-of-function (LOF) status was derived from VEP consequence terms, selected transcript impact, curated DNA/protein annotations, and structural event class. Variants were flagged as LOF when they involved stop-gain, frameshift, canonical splice donor or

acceptor consequences, transcript ablation, high-impact VEP annotations, protein-level truncating descriptions, or structural events such as deletions, duplications, insertions, and exon-level copy-number changes.

Genetic case resolution framework

Genetic resolution was performed using a hierarchical, rule-based framework that integrated inheritance models, variant pathogenicity, allelic configuration, zygosity, sex for X-linked genes, and allele-frequency information. The framework comprised 17 predefined resolution rules, encompassing canonical Mendelian inheritance patterns (autosomal recessive, autosomal dominant, and X-linked), mitochondrial inheritance, dual AR/AD inheritance genes, and frequency-based extensions applied to cases not resolved by canonical criteria.

All rules were evaluated at the gene level, considering the complete set of variants identified in each gene for a given patient. Each rule captures a distinct genetic configuration defined by the mode of inheritance, the number and type of variants, evidence of pathogenicity, loss-of-function status, zygosity, sex for X-linked genes, allele frequency, and consistency with the clinical phenotype. Rules 1–15 correspond to canonical genetic models, including autosomal recessive, autosomal dominant, X-linked, autosomal recessive/autosomal dominant, and mitochondrial inheritance. Rules 16–17 represent frequency-based extensions applied only to cases that did not meet stronger pathogenicity- or LOF-based criteria under Rules 1–15.

When more than one gene fulfilled resolution criteria within a single case, a single primary rule was assigned using a predefined priority hierarchy that placed complete biallelic or clearly dominant/X-linked/mitochondrial configurations ahead of partial, carrier-like, and frequency-based evidence. Mitochondrial variants were evaluated using dedicated criteria that reflect non-Mendelian inheritance and established pathogenicity. Cases were classified as resolved when the primary rule fulfilled complete genetic criteria, or as partially resolved when at least one plausible disease-associated variant was identified but the allelic configuration, inheritance context, or evidence level remained incomplete. Partially resolved cases were considered biologically informative and were analysed separately from fully resolved cases. Complete rule descriptions and case counts per rule are provided in Supplementary Table S4 and the accompanying Supplementary Information.

Statistical Analysis

Analyses were performed using the appropriate unit of analysis for each question. Patient-level analyses counted each individual once, whereas gene-, genotype-, and variant-level analyses allowed the same patient to contribute multiple entries when more than one gene or variant was involved. Complex genotypic configurations, including multiallelic states beyond simple monoallelic or biallelic models, were retained to reflect the genetic heterogeneity observed in IRDs. Unless otherwise specified, counts and percentages were calculated according to the relevant unit of analysis and are indicated in the corresponding figures, tables, and legends.

Data from SARA and from the SPG et al. comparison dataset were harmonised prior to analysis to ensure consistency in patient identifiers, gene symbols, clinical phenotypes, inheritance categories, and variant annotations. Clinical phenotypes were mapped to a unified phenotype dictionary based on SARA nomenclature while preserving original source annotations. Descriptive statistics were performed using R (version ≥ 4.2.0). Data manipulation, aggregation, visualisation, figure generation, and manuscript rendering were performed in R/RStudio using tidyverse, ggplot2, Quarto, and supporting packages.

OpenAI Codex was used as an AI-assisted programming tool to support code development, script debugging, and reproducibility checks associated with the analyses presented in this study. All analytical decisions, data interpretation, and final manuscript content were reviewed and curated by the authors.

Additional analytical details, including the full definition of the genetic resolution rules, are provided in the Supplementary Information.

Declarations

Funding

No specific funding was received for this work.

Ethics Approval and Consent to Participate

The SARA database is a longitudinal, observational, clinical-genomic resource comprising patients who voluntarily enrolled with the Stargardt APNES–Retina Argentina Association and provided written informed consent for genetic testing, data storage, and the use of fully anonymised clinical and molecular information for research and scientific publication. The database is institutionally maintained and formally registered with the Ministry of Justice of the Argentine Republic (Registration No. 71776). The retrospective analysis presented in this study was reviewed and approved by the Institutional Review Board (IRB) of Fundación Actyon (Approval date: 16 October 2025). Written informed consent was

obtained from all participants or their legal guardians. All procedures were conducted in accordance with Argentine Law 25,326 on the protection of personal data and with the principles of the Declaration of Helsinki and its subsequent revisions.

In addition, a subset of patients with Leber hereditary optic neuropathy (LHON) who were evaluated for participation in an international ND4 gene therapy clinical trial (ClinicalTrials.gov Identifier: NCT03153293; start date December 17, 2017, and finalisation date January 15, 2025) underwent independent ethical and regulatory review by the Advisory Commission on Regenerative Medicine and Cellular Therapies of the Argentine Ministry of Science, Technology and Innovation (MINCyT). Their inclusion in the clinical trial was approved following evaluation of the scientific, ethical, and regulatory framework governing the study.

Competing Interests

The authors declare no competing financial interests. Fundación Actyon is a non-profit organisation.

Acknowledgments

We gratefully acknowledge the patients and families for their participation, as well as the Fundación Actyon for its long-standing support of molecular diagnostic activities, patient access initiatives, and the development of the SARA clinical-genomic resource. This work is dedicated to the memory of Roser Gonzàlez-Duarte, whose scientific leadership, generosity, and commitment to research on inherited retinal disease were fundamental to this study. Her contribution and mentorship continue to inspire our work.

Author Contribution

Conceptualization, M.C., F.A., and H.D.; project administration and coordination with patient organisations, M.C. and F.A.; clinical investigation and patient phenotyping, A.A., L.E., D.H., and M.C.; genetic investigation, sample processing, bioinformatic analyses, and clinical reporting, R.V., M.T., G.M., D.H., M. S., and R.G.-D.; data curation and variant interpretation, E.F., J.M., R.V., and M.C.; formal analysis and data visualisation, E.F. and H.D.; writing—original draft preparation, M.C., F.A., and H.D; writing—review and editing, all authors.

Data Availability

All data supporting the findings of this study are available within the article and its Supplementary Information files. Additional de-identified data are available from the corresponding author upon reasonable request.

References

1. Cremers,F. P. M.,Boon,C. J. F.,Bujakowska,K.&Zeitz,C.Special Issue Introduction: Inherited Retinal Disease: Novel Candidate Genes, Genotype–Phenotype Correlations, and Inheritance Models.*Genes*9,215(2018).
2. Ben-Yosef,T.Inherited Retinal Diseases.*International Journal of Molecular Sciences*23,13467(2022).
3. Ng,Q. X.*et al.*Cost-of-illness studies of inherited retinal diseases: a systematic review.*Orphanet Journal of Rare Diseases*19,93(2024).
4. Lin,Y.*et al.*Burden and characteristics of inherited retinal diseases in China.*Scientific Reports*15,32748(2025).
5. Sangermano,R.,Galdikaité-Braziené,E.&Bujakowska,K. M.Retinal Degenerative Diseases XIX, Mechanisms and Experimental Therapy.*Advances in Experimental Medicine and Biology*1415,173–182(2023).
6. Rivolta,C.*et al.*RetiGene, a comprehensive gene atlas for inherited retinal diseases.*The American Journal of Human Genetics*112,2253–2265(2025).
7. Khan,M.,Fadaie,Z.,Cornelis,S. S.,Cremers,F. P. M.&Roosing,S.Retinal Degeneration, Methods and Protocols.*Methods in Molecular Biology*1834,3–27(2018).
8. Toulis,V.*et al.*Increasing the Genetic Diagnosis Yield in Inherited Retinal Dystrophies: Assigning Pathogenicity to Novel Non-canonical Splice Site Variants.*Genes*11,378(2020).
9. Amato,A.*et al.*Gene Therapy in Inherited Retinal Diseases: An Update on Current State of the Art.*Frontiers in Medicine*8,750586(2021).
10. Hu,M. L.*et al.*Gene therapy for inherited retinal diseases: progress and possibilities.*Clinical and Experimental Optometry*104,444–454(2021).
11. Drag,S.,Dotiwala,F.&Upadhyay,A. K.Gene Therapy for Retinal Degenerative Diseases: Progress, Challenges, and Future Directions.*Investigative Ophthalmology & Visual Science*64,39(2023).
12. Jo,D. H.,Bae,S.,Kim,H. H.,Kim,J.-S.&Kim,J. H.In vivo application of base and prime editing to treat inherited retinal diseases.*Progress in Retinal and Eye Research*94,101132(2023).

13. Brar,A. S.*et al.*Gene Therapy for Inherited Retinal Diseases: From Laboratory Bench to Patient Bedside and Beyond.*Ophthalmology and Therapy*13,21–50(2024).
14. Motta,F. L.,Martin,R. P.,Filippelli-Silva,R.,Salles,M. V.&Sallum,J. M. F. Relative frequency of inherited retinal dystrophies in Brazil.*Scientific Reports*8,15939(2018).
15. Holtan,J. P.,Selmer,K. K.,Heimdal,K. R.&Bragadóttir,R.Inherited retinal disease in Norway – a characterization of current clinical and genetic knowledge.*Acta Ophthalmologica*98,286–295(2020).
16. Sharon,D.*et al.*A nationwide genetic analysis of inherited retinal diseases in Israel as assessed by the Israeli inherited retinal disease consortium (IIRDC).*Human Mutation*41,140–149(2020).
17. Chen,T.-C.*et al.*Genetic characteristics and epidemiology of inherited retinal degeneration in Taiwan.*npj Genomic Medicine*6,16(2021).
18. Perea-Romero,I.*et al.*Genetic landscape of 6089 inherited retinal dystrophies affected cases in Spain and their therapeutic and extended epidemiological implications.*Scientific Reports*11,1526(2021).
19. Villanueva-Mendoza,C.*et al.*The Genetic Landscape of Inherited Retinal Diseases in a Mexican Cohort: Genes, Mutations and Phenotypes.*Genes*12,1824(2021).
20. Karali,M.*et al.*Genetic epidemiology of inherited retinal diseases in a large patient cohort followed at a single center in Italy.*Scientific Reports*12,20815(2022).
21. Suga,A.*et al.*Genetic characterization of 1210 Japanese pedigrees with inherited retinal diseases by whole-exome sequencing.*Human Mutation*43,2251–2264(2022).
22. Peter,V. G.*et al.*The first genetic landscape of inherited retinal dystrophies in Portuguese patients identifies recurrent homozygous mutations as a frequent cause of pathogenesis.*PNAS Nexus*2,pgad043(2023).
23. Moya,R.*et al.*Four Unique Genetic Variants in Three Genes Account for 62.7% of Early-Onset Severe Retinal Dystrophy in Chile: Diagnostic and Therapeutic Consequences.*International Journal of Molecular Sciences*25,6151(2024).
24. Mena,M. D.*et al.*Clinical and Genetic Spectrum of Stargardt Disease in Argentinean Patients.*Frontiers in Genetics*12,646058(2021).
25. Schlottmann,P. G.*et al.*Nationwide genetic analysis of more than 600 families with inherited eye diseases in Argentina.*npj Genomic Medicine*8,8(2023).
26. Instituto Nacional de Estadística y Censos.Estudionacionalsobreelperfildelaspersonascondiscapacidad:Resultadosdefinitivos2018.(2018).
27. Ciccioli,M.Articulation between an argentine patient organization, the argentinian state and researchers from the university of huazhong (china) for inclusion of 10 argentine patients in ND4 gene therapy trial for leber hereditary optic neuropathy (LHON).(2019).
28. Ciccioli,M.&Antacle,A.Revisión sobre terapia génica aprobada y recuperación visual en la amaurosis congénita de leber por mutación en el gen RPE65.*Oftalmología Clínica y Experimental*17,eXX–eXX(2024).
29. Lee,W.*et al.*A genotype-phenotype correlation matrix for ABCA4 disease based on long-term prognostic outcomes.*JCI Insight*7,e156154(2021).
30. Hanany,M.,Rivolta,C.&Sharon,D.Worldwidecarrierfrequencyandgeneticprevalenceofautosomalrecessiveinheritedretinaldiseases.*Proceedings of the National Academy of Sciences*117,2710–2716(2020).

Figures

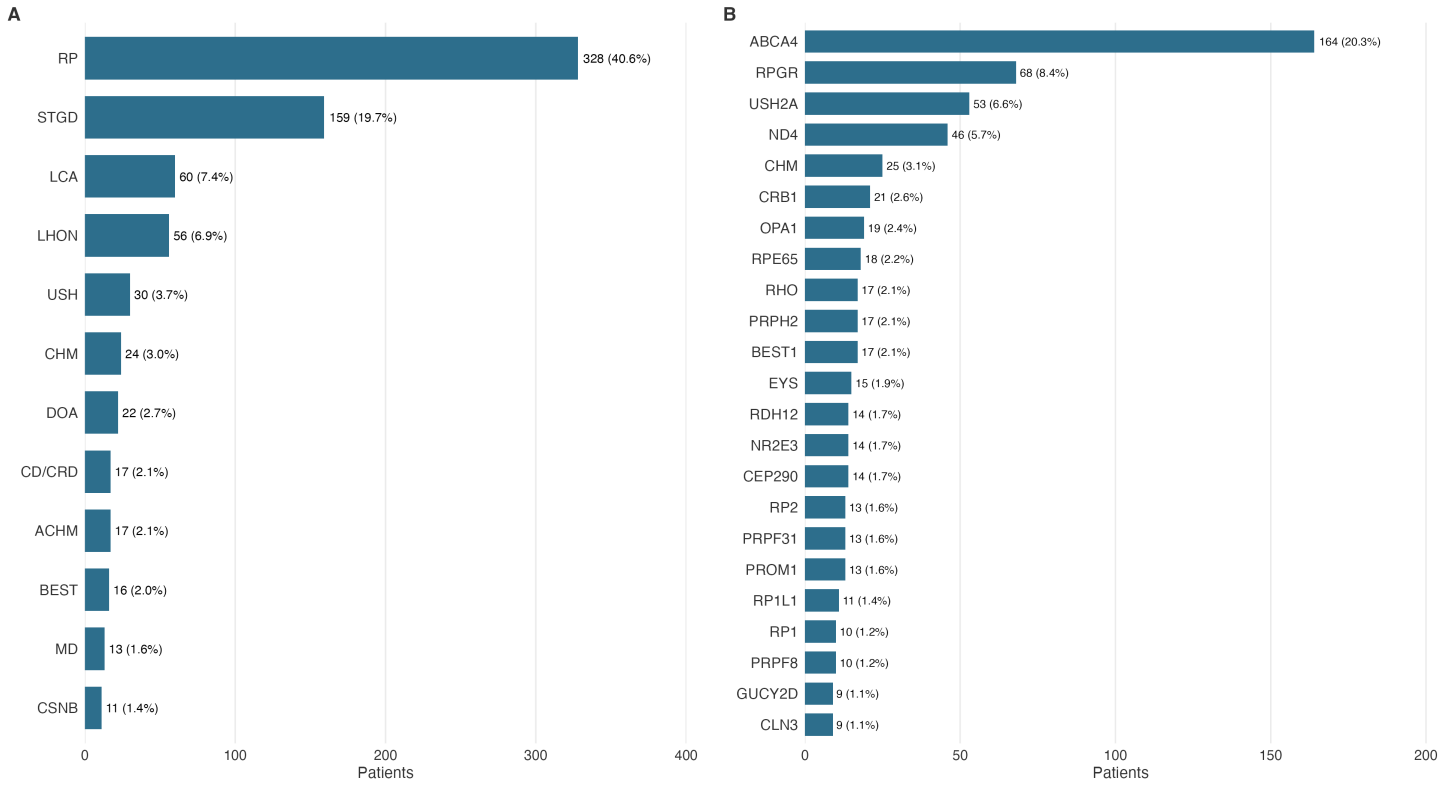


Figure 1

Clinical and gene-level architecture of the SARA cohort. (A) Distribution of inherited retinal disease (IRD) diagnoses with a prevalence at or above 1% in the cohort. Bars indicate the number of patients, with percentages calculated among genetically solved cases. (B) Recurrently implicated genes with frequencies at or above 1% among genetically solved patients, calculated at the patient–gene level. Patients carrying variants in multiple IRD-associated genes could contribute to more than one gene category. Genes with frequencies below 1% are reported in Supplementary Table S2.

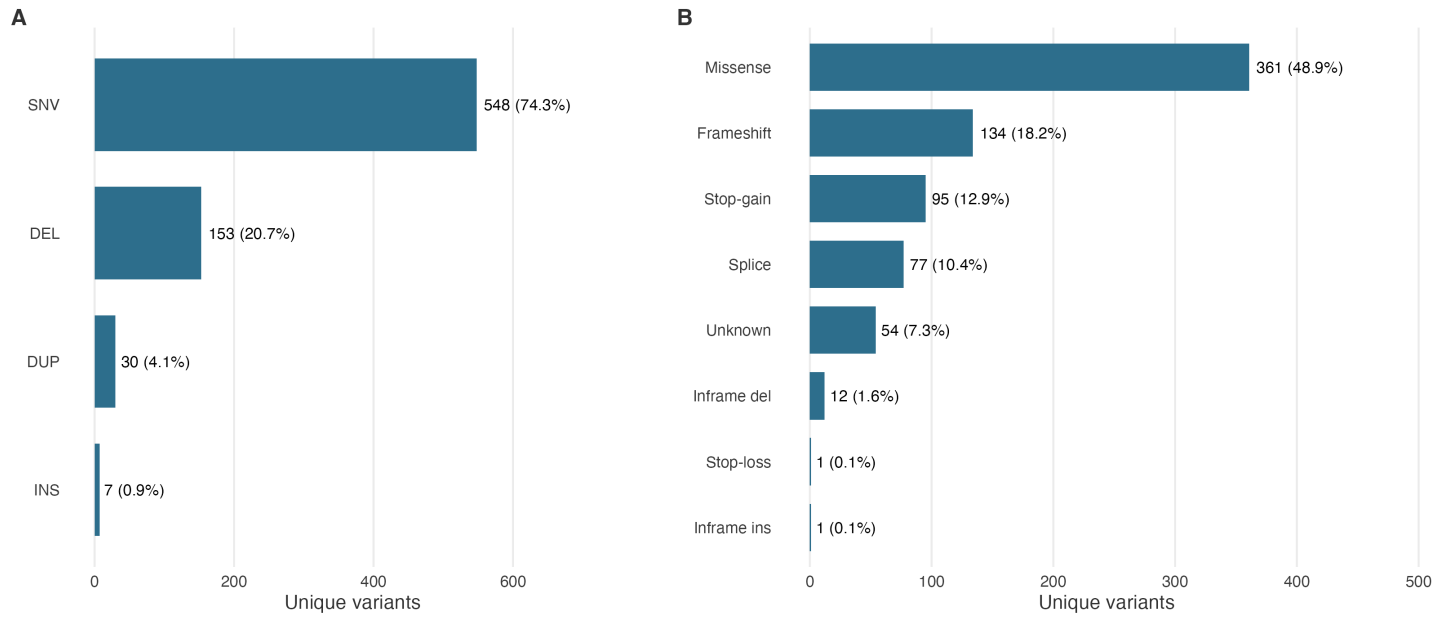


Figure 2

Genotype–inheritance patterns in the SARA cohort. Stacked bars represent 831 patient–gene configurations assigned to inheritance-compatible genotype patterns. Individual patients may contribute more than one configuration when variants were identified in multiple IRD-associated genes, including secondary or non-primary findings.

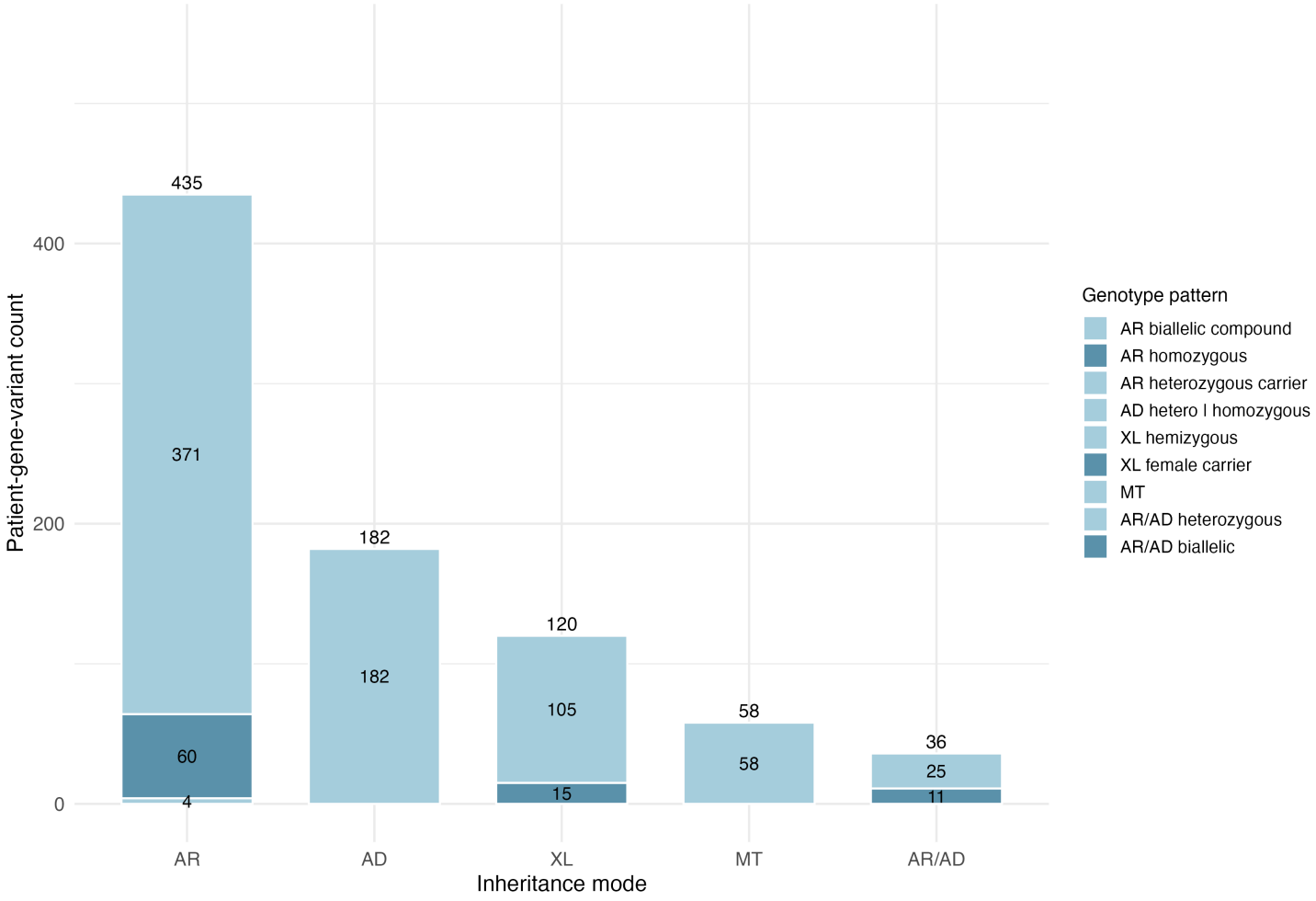


Figure 3

Distribution of DNA variant types and predicted protein-level consequences across unique variants identified in the SARA cohort. SNVs and missense changes predominated, with a substantial contribution from loss-of-function variants, including frameshift and stop-gain alterations. Variants lacking an annotated protein-level consequence were classified as Unknown. Synonymous variants (n = 3) were retained in the source tables but omitted from the protein-level display.

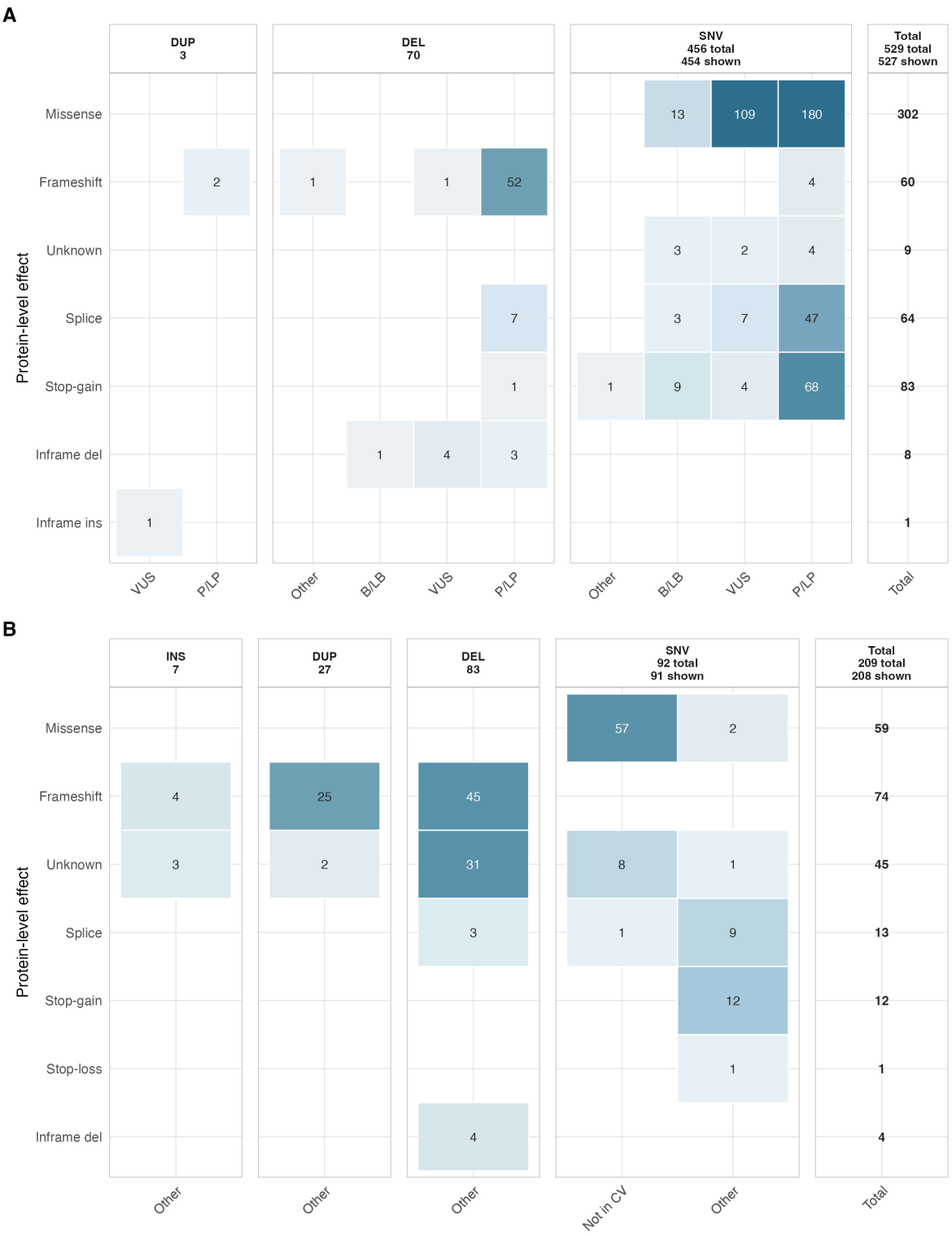


Figure 4

Heatmap representation of the molecular annotation landscape of reported and novel variants. Variants are stratified by DNA variant type, protein-level effect, and ClinVar classification. Each cell corresponds to a distinct variant category, with values indicating counts of unique variants rather than patient-level observations. Facet headers report full DNA-class totals; when synonymous variants were omitted from the protein-level display, the number of displayed non-synonymous variants is shown in parentheses. Right-hand marginal totals indicate the number of displayed variants within each protein-level effect category. Synonymous variants were retained in the source tables but omitted from this protein-level display. Darker shading denotes higher variant counts.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SARAsuppmat.pdf