AR GENE VARIANTS: A CURATED DATABASE AND CLINICAL CORRELATIONS

Reginaldo José Petrolli[1,4]; Raquel Ramos [2]; Gustavo Silveira Soares[3]; Clara Bruna Fernandes da Silva[1]; Sorahia Domenice[2]; Berenice Mendonca[2]; Rafael Batista[2].

[1]Programa de Pós-Graduação em Ciências Médicas, Faculdade de Medicina, Universidade Federal de Alagoas; [2] Hospital das Clínicas da Faculdade de Medicina, Universidade de São Paulo – São Paulo, [3]Faculdade de Medicina, Universidade Federal de Alagoas; [4]Hospital Universitário Professor Alberto Antunes, Universidade Federal de Alagoas, Maceió - Alagoas.

INTRODUCTION

Androgen Insensitivity Syndrome (AIS) is caused by pathogenic variants in the androgen receptor (*AR*) gene. The *AR* has three major domains: N-terminal (NTD), DNA-binding (DBD), and ligand-binding domain (LBD). Variants in *AR* lead to complete (CAIS), partial (PAIS), or mild (MAIS) phenotypes. However, interpreting *AR* variants remains challenging due to genotype–phenotype heterogeneity. A centralized and curated resource integrating genetic and clinical data is essential to improve diagnostic accuracy and deepen our understanding of AIS. In this context, the objective of this paper is to develop a curated *AR* gene variant database and to analyze their clinical correlations, including phenotype associations, pathogenicity classifications, and mutational hotspots in AIS.

MATERIALS AND METHODS

Variant compilation:
HGMD | ClinVar | PubMed | *AR* Database

Data Curation:
Duplicate removal
Classification by variant type

Interpretation:
ACMG criteria
REVEL & AlphaMissense prediction

Population Frequency
gnomAD & ABraOM

Statistical Analysis
Burden test → AIS vs Controls (46,XY)

RESULTS

687 AR variants from 1,135 publications →

- MAIS (46)
- PAIS (221)
- CAIS (431)

- Missense (421)
- InDels (111)
- Nonsense (84)

- Pathogenic (103)
- Likely pathogenic (461)
- VUS (68)

KEY FINDINGS:

1. **LBD** = most affected domain (**369 variants**);
2. Indels:
   - In CAIS → mostly small frameshifts (n = 64)
   - In MAIS → nonframeshift, located in tandem repeats;
3. **Burden test:** significant enrichment of small indels + CNVs in AIS vs controls
4. Hotspots: conserved arginine residues at codons 841 and 856

⚠ 60 variants associated with more than 1 phenotype

DISCUSSION

Our findings underscore the LBD as a mutation-rich and functionally critical region of the *AR* gene, with recurrent hotspots at key arginine residues. Several LBD variants are associated with more than one AIS phenotype (most commonly CAIS and PAIS) with some mutations even linked to all three forms: CAIS, PAIS, and MAIS. This underscores the genetic and phenotypic heterogeneity of AIS and the difficulty of establishing clear genotype/phenotype correlations in this syndrome.

CONCLUSION

Our findings underscore the LBD as a mutation-rich and functionally critical region of the *AR* gene, with recurrent hotspots at key arginine residues. The database provides a high-resolution map of *AR* variants, integrating predictive, clinical, and population data. This resource enhances the interpretation of variant pathogenicity, refines genotype–phenotype associations, and supports clinical decision-making and future research in AIS.

REFERENCES

The Human Gene Mutation Database (HGMD). HGMD Search. Available at: https://www.hgmd.cf.ac.uk/ac/search.html.
National Center for Biotechnology Information (NCBI). *ClinVar*. Available at: https://www.ncbi.nlm.nih.gov/clinvar/
ANDROGENDB. The Androgen Receptor Gene Mutations Database. McGill University. Disponível em: https://androgendb.mcgill.ca/
National Center for Biotechnology Information (NCBI). *PubMed*. Disponível em: https://pubmed.ncbi.nlm.nih.gov/

AGRADECIMENTOS: