APRIMORAMENTO DE UMA BASE DE ANOTAÇÕES DE VARIANTES GENÉTICAS DE L1 EM DISTÚRBIOS DO NEURODESENVOLVIMENTO

Victor Angelo Legat Cerqueira (PIBIC/FA), Adriano Ferrasa (Orientador), e-mail: ferrasa@uepg.br.

Universidade Estadual de Ponta Grossa, Departamento de Informática.

Ciência da Computação (1.03.00.00-7) / Metodologia e Técnicas da Computação (1.03.03.00-6)

Palavras-Chave: *Long interspersed element-1 (L1), SNP, base de conhecimento biológico, distúrbios do neurodesenvolvimento, transcriptoma humano.*

Resumo

Os elementos LINE-1 (L1) estão ligados a distúrbios do neurodesenvolvimento (DN). Este trabalho expandiu e aprimorou a base SNPL1DB, analisando dados de RNA-Seq de 717 amostras de pacientes. Foi desenvolvida uma nova plataforma web (Next.js) com filtros interativos e exportação de dados. A base aprimorada constitui recurso valioso para investigar a correlação entre variantes transcricionais de L1 e DN, com potencial para apoiar avanços diagnósticos e terapêuticos.

Introdução

Elementos L1 são retrotransposons cuja instabilidade está associada a DN como autismo e esquizofrenia (Singer *et al.*, 2010). Apesar da relevância de bases de dados biológicas (Baxevanis; Bateman, 2015), a SNPL1DB apresentava limitações de dados e usabilidade. Este projeto buscou expandi-la e aprimorá-la, integrando mais dados de transcriptoma e desenvolvendo interface interativa que auxilie no entendimento do papel transcricional de L1 nos DNs.

Materiais e Métodos

Foram analisadas 717 amostras de RNA-Seq de pacientes. Utilizou-se um *pipeline* de bioinformática com alinhamento via STAR (Dobin *et al.*, 2013) e chamada de variantes com GATK (Depristo *et al.*, 2011). A nova plataforma web foi construída em Next.js e JavaScript, com dados importados em PostgreSQL. Conteinerizada com Docker, implementa filtros interativos (cromossomo, gene, tipo de variante) e exportação de dados, facilitando a análise.

Resultados e Discussão

O *pipeline* gerou um catálogo de variantes de L1 das 717 amostras. A SNPL1DB aprimorada oferece plataforma web robusta, permitindo consultas com filtros multinível, visualização de metadados e exportação. A evolução transforma a base em recurso centralizado, viabilizando estudos de correlação em larga escala entre variantes transcricionais de L1 e fenótipos de DN, superando limitações anteriores e dispersão da literatura. A base está disponível em https://bioinfo.deinfo.uepg.br/snpL1/.

Conclusão ou Conclusões

A expansão e aprimoramento da SNPL1DB fornece um recurso inédito que integra dados genômicos e transcriptômicos de L1 em DNs. A combinação de um *pipeline* robusto com uma plataforma moderna resultou numa ferramenta interativa e acessível. Este trabalho consolida um recurso único com potencial para catalisar novas descobertas sobre os mecanismos moleculares dos distúrbios do neurodesenvolvimento e apoiar o desenvolvimento de futuros alvos diagnósticos e terapêuticos.

Agradecimentos

À Fundação Araucária pelo apoio financeiro, ao Prof. Dr. Adriano Ferrasa pela orientação e LCAD/UEPG por suporte em infraestrutura computacional.

BAXEVANIS, A. D.; BATEMAN, A. The importance of biological databases in biological discovery. **Current Protocols in Bioinformatics**, [s. l.], v. 1.1, n. 50, p. 1.1.1-1.1.8, 2015.

DEPRISTO, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. **Nature Genetics**, [s. l.], v. 43, n. 5, p. 491–501, 2011.

DOBIN, A. *et al.* STAR: ultrafast universal RNA-seq aligner. **Bioinformatics (Oxford, England)**, [s. l.], v. 29, n. 1, p. 15–21, 2013.

SINGER, T. *et al.* LINE-1 retrotransposons: Mediators of somatic variation in neuronal genomes?. **Trends in Neurosciences**, [s. l.], v. 33, n. 8, p. 345–354, 2010.

ISSN 1676-0093