

SOFTWARE EPHLA-CONVERTER - UM WEBSITE PARA AUTOMAÇÃO DA INFERÊNCIA DO HLA DE ALTA RESOLUÇÃO.

José Renato Pereira de Sousa Barroso (bolsista PIBIT/UFPI), Herton Luiz Alves Sales Filho (colaborador, bolsita PIBIC/UFPI), Semiramis Jamil Hadad do Monte (co-orientadora, Depto. de Parasitologia e Microbiologia), Luiz Cláudio Demes de Mata Sousa (orientador, Depto. de Informática e Estatística - UFPI).

Introdução

Definir HLA em alta resolução é importante para transplante de medula óssea entre indivíduos não aparentados e para análise de compatibilidade em pacientes sensibilizados aguardando transplante de órgãos sólidos facilitando identificação dos “mismatches” aceitáveis, utilizando o algoritmo HLAMatchmaker, recentemente automatizado no software EpHLA (Sousa, Sales Filho et al., 2011).

Contudo, custos elevados e falta de acesso ao método impedem que a tipificação HLA de alta resolução seja rotineiramente aplicada a todos os pacientes sensibilizados contra antígenos HLA. Usualmente, técnicas de média resolução geram tipagens que definem apenas grupo alélico (2 primeiros dígitos) e a lista de possibilidades de alelos dentro desse grupo (código NMDP), onde o alelo do paciente está contido. Utilizando dados de frequências alélicas descritos em populações de referência é possível inferir alelos HLA (alta resolução) a partir da tipagem em média resolução. No entanto, essa inferência depende de diversas consultas manuais a tabelas de frequências de populações e, tais consultas manuais dificultam ou impedem estudos de histocompatibilidade envolvendo centenas ou milhares de doadores e receptores cujas tipagens HLA estão disponíveis em média ou baixa resolução.

No presente trabalho, descreveremos o programa web EpHLA-Converter que permite inferir automaticamente tipagens HLA em alta resolução a partir de tipagens em média resolução de milhares de doadores e/ou receptores de uma determinada população. A inferência, dessas milhares de tipagens, pode ser realizada por meio de processamento em lote. Além disso, o programa é flexível o suficiente para que qualquer tabela de frequência de população possa ser acrescentada e utilizada, permitindo inferências mais acuradas para uma população específica. O programa EpHLA-Converter foi desenvolvido de forma a permitir a fácil integração de ferramentas de terceiros. Para comprovar essa facilidade, o EpHLA foi integrado ao EpHLA-Converter utilizando-se a tecnologia JSON [JSON, 2012].

Metodologia

O programa EpHLA-Converter foi desenvolvido em linguagem Python [PYTHON, 1990] e uma *framework* pública Web2py [WEB2PY, 2007]. O programa utiliza como dados de entrada as tipagens HLA em média ou baixa resolução de doadores e/ou receptores e retorna uma lista de tipagens em alta resolução. Para inferir a tipagem de alta resolução a partir da tipagem de baixa ou média resolução, o programa consulta bancos de dados públicos tais como o IMGT (ImMunoGeneTics program) e o NMDP (National Marrow Donor Program). No IMGT, o EpHLA-Converter verifica a validade de todos os alelos de alta resolução retornados (Marsh, Albert et al., 2010), bem como a

equivalência entre a alta resolução e a sorologia correspondente. No NMDP o EpHLA-Converter decodifica o código NMDP em alelos e as respectivas frequências por etnias da população americana (Maiers, Gragert et al., 2007), permitindo ao usuário escolher o melhor alelo para o caso. No processo de inferência da tipagem de alta resolução a partir de tipagens em biologia molecular (i.e. *locus* + grupo alélico + código NMDP) o usuário precisa informar os seguintes dados de entrada: (i) população de referência; (ii) locus HLA; (iii) grupo alélico e; (iv) código NMDP. Na inferência da tipagem de alta resolução a partir de tipagens em sorologia (i.e. *locus* + grupo sorológico), o usuário informa: (i) população de referência; (ii) locus HLA e; (iii) sorologia. Em ambos os casos, o programa EpHLA-Converter retorna uma lista com os alelos em alto nível ordenados por frequência.

Resultados e Discussão

Estudo 1

Utilizamos o EpHLA-Converter para inferir os alelos mais prováveis de um paciente para o locus A. A média resolução informada foi A*02:BXYG, A*02:WHG. Considerando apenas o grupo alélico (A*02), ele seria considerado homocigoto. Utilizando o segmento hispânico da população americana como referência, a tipagem HLA em alta resolução de maior frequência encontrada pelo EpHLA-Converter foi A*02:01 e A*02:02. Considerando que o código WHG contempla apenas os alelos A*02:02 / A*02:47, isso elimina a possibilidade de homocigotidade para A*02:01. De fato, a tipagem por alta resolução realizada posteriormente confirmou heterocigotidade: A*02:01 e A*02:02.

Estudo 2

Utilizamos um banco de dados composto por 5.742 doadores de plaquetas, cadastrados no Laboratório de Imunologia Plaquetária do Hemocentro-Unicamp (LIPH-Unicamp), tipados em média resolução para HLA A e B e, como população de referência, escolhemos o segmento caucasóide americano. O banco de dados foi convertido para alta resolução utilizando-se duas abordagens diferentes: Entrada 1- Forneceu-se ao EpHLA-Converter uma lista de doadores e suas respectivas tipagens HLA em média resolução; Entrada 2 - Forneceu-se ao EpHLA-Converter uma lista de doadores de plaquetas e suas respectivas tipagens HLA em baixa resolução. As listas de alelos em alta resolução retornadas em ambas abordagens foram comparadas e as discordâncias entre elas foram sumarizadas conforme mostrado na **Tabela 1**.

Tabela 1 – Resultado da conversão do banco de dados do LIPH-Unicamp.

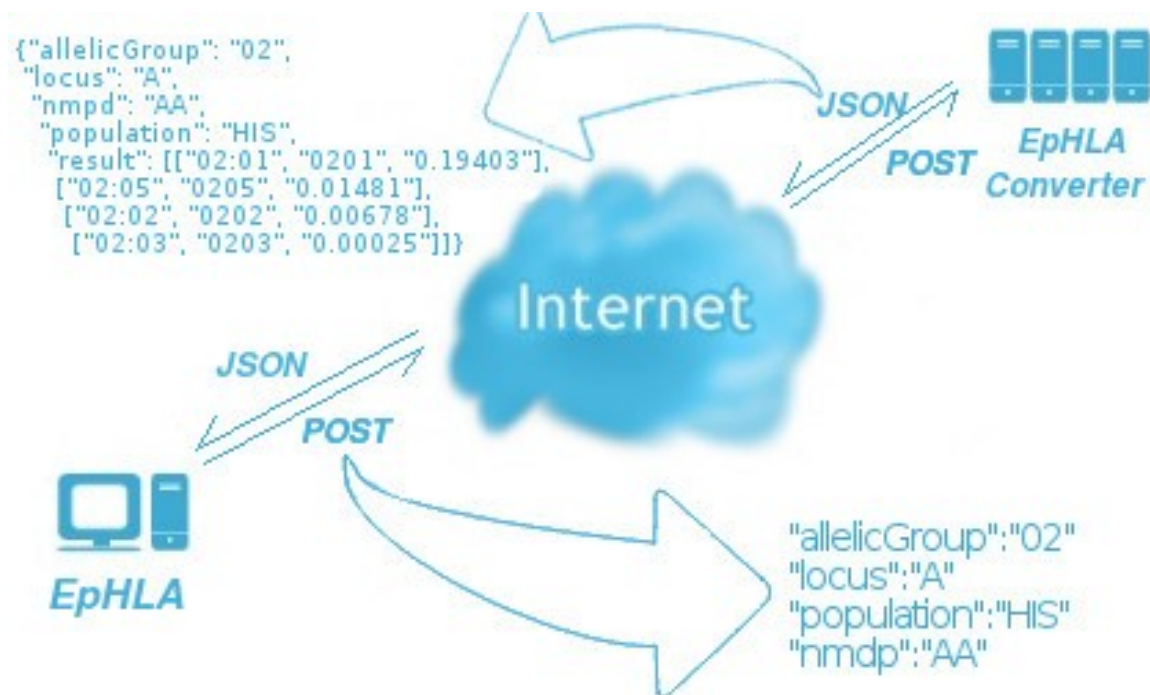
| <i>Locus</i> | Consultas realizadas | Consultas discordantes (Entrada 1 x Entrada 2) | % Discordante |
|--------------|----------------------|--|---------------|
| A | 11.484 | 1.186 | 10,33% |
| B | 11.484 | 2.892 | 25,18% |

Comparando as duas abordagens, detectou-se que 53,87% dos doadores tem pelo menos uma tipagem que é mais acuradamente inferida com o uso da média resolução.

Estudo 3

O EpHLA é uma ferramenta que permite o estudo de histocompatibilidade entre doadores e receptores utilizando o algoritmo HLAMatchmaker. e foi desenvolvida e patenteada na Universidade Federal do Piauí. O EpHLA precisa de tipagens em alta resolução para poder comparar um doador a um ou mais receptores e verificar a histocompatibilidade entre eles. No entanto, a tipagem em alta resolução é cara e, sendo assim, a abordagem utilizada é inferi-la a partir de tabelas de freqüência de população. No caso em que um doador precisa ser estudado quanto à compatibilidade com um receptor, apenas doze inferências são necessárias: seis para o receptor e seis para o doador, considerando apenas os *loci* A, B e DR. Consultando as tabelas de freqüência manualmente, o tempo é consideravelmente irrelevante. No entanto, quando utiliza-se o EpHLA para comparar um doador a dezenas, centenas ou milhares de receptores em listas de transplante, o grande número de inferências torna a tarefa praticamente inviável. Para resolver esse problemas, integrou-se o EpHLA ao EpHLA-Converter por meio das tecnologias REST [REST, 2002] e JSON. A integração é realizada conforme mostra a **Figura 1**.

Figura 1 - Integração de um ferramenta de terceiro com o EpHLA-Converter



Conclusão

O EpHLA-Converter permite inferir alelos HLA mais prováveis a partir de tipagens em baixa e média resolução, utilizando processamento em lote, facilitando estudos de histocompatibilidade envolvendo milhares de doadores ou receptores. A inferência de tipagens de alta resolução a partir de tipagens de média resolução reduz custos e permite a realização de estudos de histocompatibilidade mais acurados. Além disso, as tecnologias REST e JSON utilizadas pelo EpHLA-Converter permitem a fácil integração de ferramentas de terceiros ao programa desenvolvido.

Apoio: UFPI, Laboratório de Imunogenética e Biologia Molecular – UFPI.

Referências bibliográficas:

SOUSA, L. C. D. M. et al. EpHLA: An innovative and user-friendly software automating the HLAMatchmaker algorithm for antibody analysis. *Transpl Immunol [S.l.]*, v. 25, n. 4, p. 210-6, Dec 2011.

MARSH, S. G. et al. Nomenclature for factors of the HLA system, 2010. *Tissue Antigens [S.l.]*, v. 75, n. 4, p. 291-455, Apr 2010.

MAIERS, M. et al. High-resolution HLA alleles and haplotypes in the United States population. *Hum Immunol [S.l.]*, v. 68, n. 9, p. 779-88, Sep 2007.

JASON (2012) (JavaScript Object Notation) is a lightweight data-interchange format. Disponível em: <http://www.json.org/>

REST (2002) (representational state transfer) is an approach for getting information content from a Web site by reading a designated Web page that contains an XML (Extensible Markup Language) file that describes and includes the desired content. Disponível em: <http://searchsoa.techtarget.com/definition/REST> .

PYTHON (1990) Python is a programming language that lets you work more quickly and integrate your systems more effectively, is a remarkably powerful dynamic programming language that is used in a wide variety of application domains. Disponível em: <http://python.org/>

WEB2PY (2007) Free open source full-stack framework for rapid development of fast, scalable, secure and portable database-driven web-based applications. Written and programmable in Python. LGPLv3 License. Current version: 1.99.7 (2012-03-04 22:12:08) stable. Disponível em: <http://web2py.com>

Palavras chave: software. HLA. Transplante.EpHLA. Automação