



---

Cidade Universitária “Zeferino Vaz”, 08 de abril de 2022.

**SIGAD: Of. CEP nº 031/2022**

Prof. Dr. Lício Augusto Velloso  
Pesquisador Responsável

**REF.: DISPENSA DE APRESENTAÇÃO DE PROJETO DE PESQUISA PARA AVALIAÇÃO DO SISTEMA CEP-CONEP.**

Prezado Senhor,

Informamos que a pesquisa intitulada “**UM MODELO TOPOLÓGICO PARA A DIVERSIDADE CELULAR DE SISTEMAS BIOLÓGICOS REVELA CENTENAS DE NOVOS TIPOS DE LINFÓCITOS T CD4 CIRCULANDO NO SANGUE E NO LÍQUIDO CEFALORRAQUIDIANO**”, para fins de publicação em evento científico da Faculdade de Ciências Médicas da Unicamp, cuja equipe de pesquisa é composta pelo pesquisador supracitado e por Davi Sidarta-Oliveira, propõe Descrever geometricamente o conceito de tipo celular, e usar suas propriedades topológicas para desenvolver uma nova geração de ferramentas de inteligência artificial para o estudo de sistemas biológicos e doenças humanas.

Deste modo, baseados no resumo anexado ao documento, o referido estudo não necessita tramitar pelo Comitê de Ética em Pesquisas envolvendo Seres Humanos, tendo em vista que foi utilizado dados de amostras celulares publicamente disponíveis.

Ressaltamos que se houver qualquer alteração no escopo do projeto, na qual envolva seres humanos, o CEP/Unicamp deve ser informado para fins de deliberação sobre essas mudanças.

Atenciosamente,

**Dra. Renata Maria dos Santos Celeghini  
COORDENADORA DO COMITÊ DE ÉTICA EM PESQUISA  
UNICAMP**



## **Um modelo topológico para a diversidade celular de sistemas biológicos revela centenas de novos tipos de linfócitos T CD4 circulando no sangue e no líquido cefalorraquidiano**

*Davi Sidarta-Oliveira \*, Lício Augusto Velloso*

### **RESUMO**

Tecnologias com resolução celular revolucionaram as ciências da vida e a medicina de precisão, permitindo a compreensão da composição celular de sistemas biológicos e o estudo de alta resolução de qualquer moléstia. Neste trabalho, criamos um arcabouço teórico para o estudo desta modalidade experimental que define matematicamente tipos celulares e linhagens de desenvolvimento celular, e desenvolvemos TopOMetry, uma ferramenta computacional que permite extrair tal informação de ensaios com resolução celular. Empregamos nosso método para demonstrar a existência de mais de cem tipos novos de linfócitos T CD4 circulantes no sangue humano e no líquido cefalorraquidiano ao explorar dados públicos, e mostramos que tal diversidade se manteve oculta devido a falhas conceituais em análises tidas hoje como padrão ouro. Tal diversidade também é presente no sangue de pacientes com dengue, lupus eritematoso sistêmico e esclerose múltipla. O modelo é capaz de representar mecanismos biológicos como ciclo celular, expansão clonal e diferenciação celular de maneira gráfica e intuitiva, permitindo interpretação e geração de hipóteses.

### **INTRODUÇÃO**

A recente explosão de tecnologias com resolução celular e a expectativa de sua aplicação clínica na próxima década tornou a análise de dados com muitas dimensões um desafio central para as ciências da vida e a medicina de precisão. Para interpretar e visualizar tais resultados, é necessário utilizar redução dimensional, de maneira a extrair as informações relevantes de grandes corpos de dados e representá-las graficamente. Em termos teóricos, tal tarefa tem uma associação direta com o conceito de identidade celular e com o cenário epigenético de Waddington, tornando a inferência de trajetórias de diferenciação algo intuitivo e natural. Atualmente, empregam-se modelos que assumem diversas características sobre esse cenário, o que enviesar resultados e levar a conclusões precipitadas. Para se recuperar o máximo possível de informação biológica, é necessário utilizar algoritmos que não assumem características sobre a estrutura geométrica dos dados, de maneira a definir matematicamente o conceito de tipo celular.

### **OBJETIVO**

Descrever geometricamente o conceito de tipo celular, e usar suas propriedades topológicas para desenvolver uma nova geração de ferramentas de inteligência artificial para o estudo de sistemas biológicos e doenças humanas.

### **MÉTODOS**

Modelamos as múltiplas linhagens celulares distintas que possam coexistir em uma mesma amostra como um cenário epigenético de Waddington expandido. Mostramos como geometricamente tal cenário pode ser perfeitamente descrito pelo Operador de Laplace Beltrami (OLB), que por construção recupera todas as informações topológicas de um conjunto



de dados quaisquer. O OLB assume apenas que há amostras o bastante para se utilizar cálculo diferencial. Utilizamos a linguagem *python* de programação para desenvolver TopOMetry, uma ferramenta modular e flexível que emprega diversos algoritmos que aproximam o OLB. O método foi validado em dados públicos e anonimizados de *single-cell RNA-seq* de células circulantes no sangue e líquido cerebrospinal de doadores humanos saudáveis, com dengue, lupus eritematoso sistêmico e esclerose múltipla. Foram usados apenas dados publicamente disponíveis e associados a comitês de ética em pesquisa.

## RESULTADOS

Estabelecemos um conceito geométrico de identidade celular baseado em informação epigenética sobre cada célula. Para explorá-lo, oferecemos uma ferramenta amigável, gratuita e universal para a análise de dados com resolução celular, que além deste propósito primário, também pode ser utilizado de modo geral para a análise de dados com muitas dimensões, como textos, imagens e grafos. Paralelamente, apresentamos uma diversidade nunca antes vista de linfócitos T CD4 na circulação humana em indivíduos saudáveis e em pacientes com dengue, lupus eritematoso sistêmico e esclerose múltipla. Esta diversidade também pôde ser encontrada no líquido cefalorraquidiano de doadores saudáveis e com esclerose múltipla. Também demonstramos a conexão natural de nosso modelo com o ciclo celular utilizando dados de medula óssea humana, e com o processo de diferenciação celular utilizando dados da embriogênese do camundongo, descobrindo cerca de 380 novas linhagens celulares.

## CONCLUSÕES

Fornecemos plataforma teórica e computacional para a descrição detalhada da diversidade celular em sistemas biológicos, para isto definindo geometricamente tipos e linhagens celulares. Com esta base sólida, foi possível descobrir mais de cem novos tipos de linfócitos T CD4 que se mantiveram ocultos até então, e definir que populações celulares previamente consideradas canônicas na verdade correspondem apenas a distorções lineares nos dados. Assim, a vasta literatura de resultados obtidos com os padrões atuais devem ser cuidadosamente revisitados, assim como significativa parte da literatura envolvendo linfócitos T CD4 ou processos de diferenciação celular.

---

Documento assinado eletronicamente por **Renata Maria dos Santos Celeghini, COORDENADORA DO COMITÊ DE ÉTICA EM PESQUISA DA UNICAMP**, em 08/04/2022, às 16:18 horas, conforme Art. 10 § 2º da MP 2.200/2001 e Art. 1º da Resolução GR 54/2017.

---



A autenticidade do documento pode ser conferida no site:  
[sigad.unicamp.br/verifica](http://sigad.unicamp.br/verifica), informando o código verificador:  
**A247DD14 DCAC4957 8E7C20C9 D28E6FÉ6**

