

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE FÍSICA DE SÃO CARLOS

# Análise Estatística da Teoria de Quase-espécies de Evolução Molecular

*Domingos Alves*

*Tese apresentada ao Instituto de Física de  
São Carlos, Universidade de São Paulo, para  
obtenção do Título de Doutor em Ciências:  
Física Básica*

USP/IFSC/SBI



8-2-001255

ORIENTADOR: *Prof. Dr. José Fernando Fontanari*

SÃO CARLOS

1999

IFSC-USP  
BIBLIOTECA B

Alves, D.

Análise Estatística da Teoria de Quase-espécies de Evolução Molecular/Domingos Alves – São Carlos, 1999.

138 p.

Tese (Doutorado)–Instituto de Física de São Carlos, 1999.

Orientador: Prof. Dr. José Fernando Fontanari

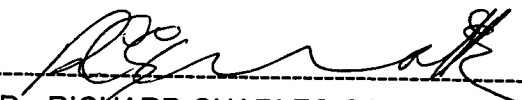
1. Quase-espécie. 2. Evolução Molecular. I. Título.




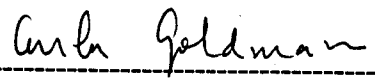
MEMBROS DA COMISSÃO JULGADORA DA TESE DE DOUTORADO  
DE DOMINGOS ALVES, APRESENTADA AO INSTITUTO DE FÍSICA DE SÃO  
CARLOS, UNIVERSIDADE DE SÃO PAULO, EM 24/03/1999.

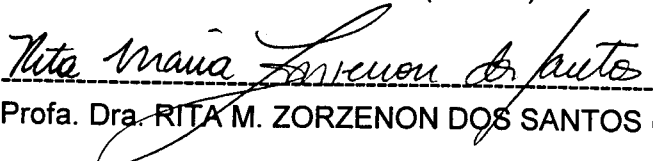
**COMISSÃO JULGADORA:**

  
-----  
Prof. Dr. JOSÉ FERNANDO FONTANARI (*Orientador*) (IFSC-USP)

  
-----  
Prof. Dr. RICHARD CHARLES GARRATT (IFSC-USP)

  
-----  
Prof. Dr. ROBERTO NICOLAU ONODY (IFSC-USP)

  
-----  
Profa. Dra. CARLA GOLDMAN (IF/USP)

  
-----  
Profa. Dra. RITA M. ZORZENON DOS SANTOS (UFF/RJ)

*Dedicado ao meu pai Domingos e às mulheres da minha  
vida: Néria, Magali e Bárbara.*

## Agradecimentos

É claro que durante a elaboração de uma tese ficamos em débito com muitas pessoas que, direta ou indiretamente nos auxiliaram, às vezes com um simples bate-papo em um insuspeitável cafezinho. Dentre essas pessoas tenho que destacar algumas que são as fundamentais.

Ao professor Fontanari, mais que um agradecimento, uma dívida de dedicar aos meus possíveis futuros alunos o mesmo profissionalismo e integridade científica que me foram dispensados. Sem sombra de dúvida sua orientação merece um destaque especial no que diz respeito à dedicação. Entretanto, gostaria de fazer menção à sua criatividade e principalmente à sua capacidade de discernir o que é relevante como produção científica.

Agradeço também aos companheiros do grupo, Daniela, Viviane, Ana e Fernando, pela ajuda, convivência amiga e conversa fácil. Particularmente, ao Camilo, ao Osame e ao Paulo pelos debates e questionamentos sempre intrigantes e entusiasmados.

Gostaria de lembrar minha colaboração com o Professor Pedro Marijuan. Suas idéias sobre sistemas biológicos sempre me encantaram e influenciaram, tendo um papel importante em minha formação.

Conhecer e interagir com a professora Rita Zorzenon nestes últimos anos foi extremamente frutífero para me decidir sobre o estudo fascinante de populações virais e suas estratégias para *enganar* o sistema imunológico. A esse respeito também, tenho que agradecer ao professor Paolo Zanotto por um simples e inesquecível seminário sobre evolução viral. Realmente seu entusiasmo científico é contagiante.

Gostaria de agradecer também ao meu irmão Gustavo pela preocupação e pelo apoio. Sinto-me em débito com ele em muitos aspectos, enfim, por todos esses anos como meu irmão mais novo. Ainda com respeito à família - a minha família - Magali e Bárbara, são realmente o começo, o meio e o fim de tudo o que me move e, em particular, as verdadeiras heroínas deste trabalho de tese, enfrentando de tudo, inclusive o meu antológico mau-humor.

Finalmente, agradeço ao Conselho Nacional de Pesquisa e Desenvolvimento Tecnológico (CNPq) que permitiu a realização deste trabalho ao conceder-me uma bolsa de Doutorado.

## Lista de Símbolos

- $A_i$  → taxa de replicação das moléculas do tipo  $i$
- $A_{P^i P^j}$  → taxa de replicação de indivíduos diplóides
- $a$  → valor seletivo da seqüência mestra
- $a_0$  → valor seletivo das seqüências mutantes em um relevo de um pico
- $\alpha$  → ordem ou intensidade da interação epistática
- $\alpha_{\max}$  → valor crítico da intensidade da interação epistática acima da qual não ocorre o limiar de erro
- $d(i, j)$  → distância de Hamming entre a seqüência  $j$  e a seqüência  $i$
- $D_i$  → taxa de desagregação da molécula do tipo  $i$
- $d$  → número de zeros em uma dada seqüência
- $\bar{d}$  → distância de Hamming média normalizada de toda a população
- $\Phi_0$  → fluxo de diluição global das macro-moléculas dentro do reator
- $\Phi_k$  → fluxo de monômeros
- $\gamma$  → parâmetro que determina o quão suave é o relevo replicativo com seleção truncada
- $\gamma_{\max}$  → valor crítico de  $\gamma$  abaixo do qual não há o fenômeno do limiar de erro
- $h$  → coeficiente de dominância
- $h_{\max}$  → valor crítico do coeficiente de dominância acima do qual não há limiar de erro
- $I_i$  → molécula do tipo  $i$  dentre as  $\kappa^L$  possíveis
- $\kappa$  → número de monômeros (ou alelos) diferentes possíveis em uma posição de uma dada seqüência (ou locus)
- $L$  → comprimento de cada seqüência ou genoma
- $L_{\max}$  → comprimento máximo de uma seqüência permitido pelo limiar de erro
- $\lambda_{\max}$  → maior auto-valor da matriz de replicação
- $M_{PR}$  → elemento da matriz de mutação da seqüência com  $R$  1's para uma seqüência com  $P$  1's
- $\mu = 1 - q$  → taxa de mutação por monômero
- $\mu_c = 1 - q_c$  → taxa de mutação crítica para manter a seqüência com maior valor seletivo na população (define a localização do limiar de erro no relevo de

um pico)

$N_{ES}$  → número de seqüências possíveis no espaço de seqüências

$N$  → número de seqüências ou genomas

$N_{\min}$  → tamanho da população abaixo do qual o limiar de erro não ocorre

$\mathbf{n}$  → número de ocupação

$n_P(t)$  → número de seqüências com  $P$  1's

$P$  → número de 1 em uma dada seqüência

$p(t)$  → freqüência de monômeros do tipo 1 na população

$p^*$  → ponto fixo para a freqüência de monômeros

$\Pi_P(t)$  → freqüência das seqüências da classe  $P$  na geração  $t$

$\Pi_{P^i P^j}(t)$  → freqüência dos indivíduos diplóides

$q$  → fidelidade de replicação de cada monômero

$\rho$  → parâmetro que mede o tamanho da região chata em um relevo replicativo

com seleção truncada

$s$  → valor seletivo de um gen deletério

$s_\beta^i$  → variável indicando o monômero (ou alelo) de uma molécula do tipo  $i$  na posição  $\beta$

$\sigma_d$  → dispersão de  $d$  em torno de  $\bar{d}$

$\sigma_N$  → dispersão que leva em consideração as flutuações em  $\Pi_P$

$\sigma_N(L)$  → espalhamento da seqüência ótima na população

$t$  → tempo ou geração

$x_i = [I_i]$  → concentração da molécula do tipo  $i$

$\mathbf{W}$  → matriz de replicação

$W_{ij}$  → elemento da matriz de replicação com erro da seqüência  $j$  à seqüência  $i$

$W_P(\mathbf{n})$  → taxa de replicação relativa da seqüência da classe  $P$

$\bar{W}$  → *fitness* médio da população

$y_i = x_i / \sum_j x_j$  → concentração relativa da seqüência do tipo  $i$

## Resumo

Nesta tese propomos e estudamos um modelo alternativo para investigar a evolução de quase-espécies moleculares, no qual supomos que a população seja uma combinação aleatória das moléculas constituintes em cada geração. Essa aleatoriedade deve-se a inclusão de um procedimento adicional de amostragem da população além dos procedimentos usuais de mutação e reprodução diferenciada. O modelo, denominado modelo de amostragens, é baseado em um algoritmo que *mimetiza* procedimentos experimentais que usam técnicas de transferências em série para reproduzir o processo de evolução de microorganismos *in vitro*. Além do modelo reproduzir a solução exata do estado estacionário do modelo de quase-espécies no regime determinístico, ele permite o estudo da evolução da quase-espécie molecular em todo espaço de parâmetros de controle, incluindo o caso em que a população é finita. A generalização dessa formulação alternativa para uma classe geral de relevos de replicação permite-nos realizar um estudo bastante completo do fenômeno do limiar de erro, levando-nos a uma análise crítica sobre a generalidade desse fenômeno.



## Abstract

In this thesis we propose and study an alternative model to investigate the evolution of a molecular quasispecies, in which we assume that the population is a random combination of the constituent molecules in each generation. This randomness is due to the inclusion of an additional sampling procedure of the population, besides the usual procedures of mutation and differential reproduction. This model, termed sampling model, is based on an algorithm that *mimics* experimental procedures using serial transfer techniques to study the microbial evolutionary process *in vitro*. Besides yielding the exact steady-state solution of the quasispecies model in the deterministic limit, the sampling model allows the study of the molecular evolution in the full space of the control parameter, including the case where of population is finite. The generalization of this alternative formulation to a general class of fitness landscapes, allows us to investigate thoroughly the error threshold phenomenon, leading us to discuss critically the generality of this phenomenon in molecular evolution.

# Conteúdo

<b>Agradecimentos</b>	<b>ii</b>
<b>Lista de Símbolos</b>	<b>iii</b>
<b>Resumo</b>	<b>v</b>
<b>Abstract</b>	<b>vi</b>
<b>Conteúdo</b>	<b>vii</b>
<b>Introdução</b>	<b>1</b>
<b>1 Micro-evolução darwiniana</b>	<b>5</b>
1.1 Um sistema modelo para a evolução molecular . . . . .	6
1.2 O ansatz da cinética química . . . . .	9
1.3 O que é uma quase-espécie? . . . . .	13
1.4 O paradoxo de Eigen . . . . .	16
1.5 Análise crítica . . . . .	21
<b>2 O modelo de amostragens</b>	<b>23</b>
2.1 Evolução <i>in vitro</i> . . . . .	24
2.2 Construção de um algoritmo para a estratégia experimental . . .	27
2.3 Resultados analíticos . . . . .	32
2.4 Comparação com a formulação de Eigen . . . . .	36

## Gracias por visitar este Libro Electrónico

Puedes leer la versión completa de este libro electrónico en diferentes formatos:

- HTML(Gratis / Disponible a todos los usuarios)
- PDF / TXT(Disponible a miembros V.I.P. Los miembros con una membresía básica pueden acceder hasta 5 libros electrónicos en formato PDF/TXT durante el mes.)
- Epub y Mobipocket (Exclusivos para miembros V.I.P.)

Para descargar este libro completo, tan solo seleccione el formato deseado, abajo:

