UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA

VIVIANE MATOS GALVÃO

MODELAGEM COMPUTACIONAL DE SISTEMAS BIOLÓGICOS

Feira de Santana, BA
2009

VIVIANE MATOS GALVÃO

MODELAGEM COMPUTACIONAL DE SISTEMAS BIOLÓGICOS

Tese apresentada ao Programa de Pós-graduação em Biotecnologia, da Universidade Estadual de Feira de Santana como requisito parcial para obtenção do título de Doutor em Biotecnologia.

Orientador: Prof. Dr. José Garcia Vivas Miranda

Comissão examinadora:

Prof. Dr. José Soares de Andrade Júnior (IF-UFC)

Prof Dr. Mauro Copelli Lopes da Silva (IF-UFPE)

Prof Dr. Roberto Fernandes Silva Andrade (IF-UFBA)

Prof. Dr. Aristóteles Góes Neto (Bio-UEFS)

Prof. Dr. José Garcia Vivas Miranda (IF-UFBA)

Feira de Santana, BA
2009

Dedico esta tese a todos os bolsistas, especialmente aos da FAPESB, CNPq e CAPES.

AGRADECIMENTOS

Agradeço ao professor e orientador José Garcia Vivas Miranda, pelo acompanhamento, discussão e revisão dos trabalhos. Aos colegas e colaboradores que participaram em algumas publicações que compõem esta tese.

Agradeço aos professores e colegas do grupo de Física Estatística e Sistemas Complexos e aos colegas do Programa de Pós-Graduação em Biotecnologia, que sempre fizeram do nosso ambiente um lugar interessante e propício para o amadurecimento profissional. Ressalto também a importância de uma convivência amigável com os professores e funcionários deste Programa.

Agradeço a Roberto Rivelino de Melo Moreno pelas sugestões e revisões aos trabalhos desenvolvidos nesta tese.

Finalmente, agradeço a FAPESB, pelo financiamento e a SBF, pelo apoio financeiro parcial nas participações em congressos.

RESUMO

Neste trabalho desenvolvemos modelos computacionais basedos em multi-agentes, redes regulares e redes complexas para reproduzir sistemas biológicos. Os sistemas simulados são a rede de diferenciação celular humana, a cardiomiopatia chagásica crônica depois do trasplante de células-tronco, a evolução da doença de Chagas e a cinética enzimática da invertase. A caracterização da rede de diferenciação celular humana é vital porque ela pode revelar as características do desenvolvimento embrionário humano usando as propriedades dinâmicas e topológicas do processo de diferenciação. Um melhor entendimento da doença de Chagas é necessário porque esta enfermidade é ainda incurável e afeta milhões de pessoas em todo o mundo. A modelagem computacional da cinética enzimática da invertase é importante porque ela pode descrever processos químicos complexos.

Palavras-chave: Rede de diferenciação celular humana. Doença de Chagas. Células-tronco. Cinética enzimática da invertase. Multi-agentes. Redes complexas.

ABSTRACT

In this work, we develope computational models based on multi-agents, regular lattice and complex networks to reproduce biological systems. The sytems simulated are the network of human cell differentiation, chronic chagasic cardiomyopathy after stem cell transplantation, evolution of Chagas' disease, and enzyme kinetics of invertase. The characterization of the network of human cell differentiation is vital because it can reveal features of the human embryonic development by using the topological and dynamical properties of the differentiation process. A better understanding of Chagas' disease is necessary because this illness is still incurable and affects millions of people worldwide. The computational modelling of enzyme kinetics of invertase is important because it can describe complex chemical process.

Keywords: Network of human cell differentiation. Chagas' disease. Stem cell. Enzymatic kinetics of invertase. Multi-agents. Complex Networks.

SUMÁRIO

Introdução geral

Os sistemas biológicos normalmente exibem características típicas de sistemas complexos. Estes sistemas podem ser descritos por um sistema composto de partes interconectadas no qual a análise das partes individuais pode não exibir o comportamento do conjunto. Este tipo de sistema é difícil de ser entendido porque as causas e efeitos não são claramente associados. O aumento do interesse em entender princípios gerais de sistemas biológicos esta relacionado com a evolução dos computadores, devido um computador mais potente poder desenvolver modelos biológicos mais complexos e de forma mais realista. Os modelos computacionais são importantes na caracterização e compreensão de sistemas biológicos, pois tentam explicar o comportamento auto-organizado e decobrir leis comuns das propriedades emergentes. Desta maneira, nesta tese desenvolvemos cinco modelos computacionais biológicos que podem ser classificados em dois tipos, segundo seus objetivos.

O primeiro tipo busca, essencialmente, a caracterização do sistema estudado a partir de índices estatísticos de forma a identificar padrões emergentes que ampliem nosso conhecimento sobre o mesmo. O segundo tipo busca a validação de hipóteses microscópicas a partir de fenômenos macroscópicos para uma melhor compreensão da dinâmica do sistema. O objetivo geral deste trabalho é mostrar que modelos computacionais são úteis para a ampliação do conhecimento sobre a cinética de sistemas biológicos. No primeiro capítulo faremos uma breve revisão sobre equações diferenciais, autômatos celulares, teoria de agentes e redes complexas. No segundo capítulo desenvolvemos um modelo bidimensional baseado em multi-agentes para a cardiomiopatia chagásica crônica depois do transplante de células-tronco e no terceiro capítulo estendemos este trabalho para um modelo tridimensional. Um modelo tridimensional baseado em agentes para a evolução da doença de Chagas é apresentado no quarto capítulo. O quinto capítulo apresenta a rede de diferenciação celular humana. No sexto capítulo desenvolvemos um modelo tridimensional baseado em agentes para a cinética enzimática da invertase. Finalmente, as conclusões gerais destes modelos computacionais biológicos são dadas no final desta tese.

CAPÍTULO 1

Modelagem de sistemas complexos

Resumo

Os sistemas complexos possuem um grande número de componentes que atuam segundo regras que podem mudar no tempo e não são bem estabelecidas. A modelagem deste tipo de sistema é muito complicada e isto está relacionado ao fato que os modelos matemáticos simplificam a descrição de sistemas reais. Atualmente, as técnicas mais usadas para modelar sistemas complexos são equações diferenciais, autômatos celulares, teoria de agentes e redes complexas. Neste texto faremos uma breve revisão sobre estas diferentes teorias.

1.1. Introdução

Um sistema complexo pode ser descrito como sendo composto de partes interconectadas em que o comportamento coletivo não é evidenciado na análise de elementos individuais. Esta característica é chamada de emergência e ela não é trivial, pois relacionar as ações individuais com o comportamento coletivo não é simples (Ottino, 2004). Geralmente, os sistemas complexos são difíceis de entender (Mitchell, 2006) e eles são altamente sensíveis às condições inicias ou a pequenas perturbações, uma vez que o número de componentes independentes interagindo é grande (Whitesides e Ismagilov, 1999). Este tipo de sistema partilha algumas propriedades como adaptação, não-linearidade e auto-organização. A adaptação é a capacidade de mudar e aprender com a experiência. A não-linearidade ocorre, pois o comportamento do conjunto não pode ser expresso como o comportamento das partes. A auto-organização é um processo no qual a organização interna do sistema aumenta de complexidade sem sofrer ações de forças externas (Bar-Yam, 1997; Ottino, 2004).

A modelagem computacional de sistemas complexos utiliza modelos simplificados, com o objetivo de descrever a evolução com um pequeno número de parâmetros (Gleria, 2004; Ioannou e

Pitsillides, 2007). Desta forma, a simulação tenta explicar o comportamento auto-organizado e descobrir leis comuns para as propriedades emergentes. O aumento do interesse em compreender os princípios gerais deste tipo de sistema está relacionado com a evolução dos computadores, pois o uso de computadores mais poderosos possibilita o desenvolvimento de modelos mais complexos (Mitchell, 2006). Exemplos de modelos computacionais para sistemas complexos incluem o mercado financeiro (Muñoz e cols., 2007; Queirós e cols., 2007), propagação de ondas (Oliveira, 2002), precipitação de chuva (Andrade e cols., 2002; Pinho e Andrade, 1998), crescimento de câncer (Galvão e Vivas, 2008; Reis e cols., 2009), regulação imune em uma rede de linfócitos (Ribeiro et al., 2007), evolução de malária (Zorzenon dos Santos e cols., 2007), desenvolvimento da infecção causada pelo HIV (Zorzenon dos Santos e Coutinho, 2001) e cinética da doença de Chagas após o transplante de células-tronco (Galvão e cols., 2008; Galvão e Vivas, 2009). Atualmente, as técnicas mais usadas para modelar sistemas complexos são equações diferenciais, autômatos celulares, teoria de agentes e redes complexas. Neste texto apresentaremos todas estas diferentes técnicas.

1.2. Equações diferenciais

Matematicamente, as equações diferenciais descrevem relações entre funções desconhecidas e suas derivadas usando uma ou várias variáveis. Este tipo de equação é muito utilizado na construção de modelos em matemática, física, engenharia e economia. O uso de equações diferenciais na modelagem possibilita que a distribuição espacial e temporal seja contínua. A solução de uma equação diferencial pode existir e não ser única ou pode não existir. As soluções constantes ocorrem quando as derivadas são nulas e são chamadas de pontos de equilíbrio (Braun, 1992; Verhulst, 2000).

1.2.1. Equações diferenciais ordinárias

Uma equação diferencial ordinária (EDO) contém somente funções de uma única variável independente e as derivadas desta variável. O comportamento de um sistema não-autônomo pode ser descrito por uma EDO que depende explicitamente do tempo da seguinte forma:

$$\frac{dx}{dt} = \dot{x} = f(x,t) \tag{1}$$

A ordem de uma equação diferencial é a ordem da mais alta derivada que está presente na equação. As condições suficientes para a existência e unicidade de uma solução da equação diferencial são definidas pelo teorema de Picard (Braun, 1992). Este teorema considera o problema de valor inicial dado pela eq. (1) e por $x(t_0) = x_0$. Se a função f e a derivada de f em função de t são contínuas na vizinhança de t_0, existe uma solução única para o problema de valor inicial.

O sistema é dito autônomo se suas derivadas não dependem explicitamente do tempo, ou seja, quando a eq. (1) não depende da variável t.

$$\dot{x} = f(x) \tag{2}$$

Uma EDO não-autônoma pode ser reduzida a uma autônoma pela introdução de mais uma dimensão no sistema. Para isto, considera-se o tempo como uma variável adimensional, ou seja, $t = r,$ tal que $dr/dt = 1$. Da mesma forma, uma EDO de segunda ordem pode ser transformada em duas equações de primeira ordem, bastando considerar uma nova variável $y = \dot{x}$ (Fiedler-Ferrara e Prado, 1994).

As EDO com atraso temporal possuem termos que dependem dos valores das variáveis em instantes de tempo anteriores ao definido pela variável de integração t. Esta dependência possibilita os modelos computacionais simularem a evolução de diversos tipos de sistemas (Murray, 2002). O acréscimo do atraso temporal na EDO pode modificar os pontos de equilíbrio e suas órbitas. Um atraso temporal pode ser adicionado a eq. (2). Então,

$$\dot{x} = f(x(t), x(t-\tau)) \tag{3}$$

onde τ é o parâmetro do retardo temporal. O acréscimo de atrasos distribuídos resulta em uma equação diferencial integral,

$$\dot{x} = f(x(t), \int_{-\infty}^{t} x(\tau) W(t-\tau) d\tau) \tag{4}$$

onde W é a função de memória e representa a influência da variável $x(\tau)$ na evolução do sistema. Os valores prévios das variáveis quando um sistema esta em equilíbrio é constante. Portanto, o modelo para um sistema em equilíbrio deve manter todos os pontos do modelo original sem atraso temporal.

1.2.2. Estabilidade de sistemas não-lineares

O comportamento de um sistema não-linear pode ser estudado através de sua linearização em torno dos pontos de equilíbrio, investigando-se localmente a estabilidade destes pontos e construindo diagramas locais, através da linearização (Fiedler-Ferrara e Prado, 1994). Considere o sistema de equações diferenciais não-lineares

$$\begin{aligned} \dot{x} &= f(x, y) \\ \dot{y} &= g(x, y) \end{aligned} \tag{5}$$

para o qual existe um ponto de equilíbrio $P^* = (x^*, y^*)$. A expansão em série de Taylor tomando somente os termos de primeira ordem é dada por,

$$\begin{aligned} \dot{x} &= f(x, y) = f(x^*, y^*) + \frac{\partial f}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*)(y - y^*) \\ \dot{y} &= g(x, y) = g(x^*, y^*) + \frac{\partial g}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial g}{\partial y}(x^*, y^*)(y - y^*) \end{aligned} \tag{6}$$

Definindo-se duas novas variáveis $\tilde{x} = x - x^*$ e $\tilde{y} = y - y^*$ e adicionalmente observando que $\dot{\tilde{x}} = \dot{x}$, $\dot{\tilde{y}} = \dot{y}$ e $f(x^*, y^*) = g(x^*, y^*) = 0$, obtemos o sistema linear,

$$\begin{aligned} \dot{\tilde{x}} &= a\tilde{x} + b\tilde{y} \\ \dot{\tilde{y}} &= c\tilde{x} + d\tilde{y} \end{aligned} \tag{7}$$

Com esta linearização, a matriz dos coeficientes é

$$J = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \partial f / \partial x & \partial f / \partial y \\ \partial g / \partial x & \partial g / \partial y \end{bmatrix} \tag{8}$$

onde J é a matriz Jacobiana calculada no ponto de equilíbrio. O sistema linearizado (eq. 5) utilizando notação vetorial é escrito da seguinte forma:

$$\dot{\tilde{x}} = J\tilde{x} \tag{9}$$

Uma solução geral para a eq. 9 é dada por

$$\widetilde{x} = \widetilde{x}_0 e^{\lambda t} \tag{10}$$

onde o termo $\widetilde{x}_0$ é um autovetor e os autovalores λ são as raízes do polinômio característico,

$$\left| J - \lambda I \right| = 0 \tag{11}$$

onde I é a matriz identidade. Esta aproximação é válida somente para pontos próximos do ponto de equilíbrio x^* e ela mostra como o sistema evolui quando as condições iniciais desviam-se do ponto de equilíbrio. Assim, a análise de estabilidade de uma solução estacionária é reduzida ao estudo das soluções correspondentes do sistema linearizado.

A estabilidade dos pontos de equilíbrio (Fiedler-Ferrara e Prado, 1994; Murray, 2002; Thompson e Stewart, 2002) é classificada de acordo com o sinal dos autovalores. Considerando o caso bidimensional existem várias possibilidades para os autovalores (λ_1, λ_2). No caso de autovalores reais tem-se: se λ_1, λ_2 apresentam o mesmo sinal e $\lambda_1, \lambda_2 < 0$ o ponto é chamado de um *nó* assintoticamente estável (Fig. 1.1.a); se λ_1, λ_2 apresentam o mesmo sinal e $\lambda_1, \lambda_2 > 0$ o ponto é um *nó* instável (Fig. 1.1.b); se λ_1, λ_2 apresentam sinais diferentes o ponto é um *ponto de sela hiperbólico* instável (Fig. 1.1.c); se $\lambda_1 = \lambda_2 > 0$ o ponto é um *nó impróprio* instável (Fig. 1.1.d); se $\lambda_1 = \lambda_2 < 0$ o ponto é um *nó impróprio* assintoticamente estável (Fig. 1.1.e). No caso de autovalores complexos conjugados considera-se o autovalor dado por $\lambda = \alpha \pm i\beta$ com $\beta \neq 0$, onde a estabilidade do ponto de equilíbrio é dada por α: se $\alpha > 0$ o ponto é um *foco* instável (Fig. 1.1.f); se $\alpha < 0$ o ponto é um *foco* assintoticamente estável (Fig. 1.1.g); se $\alpha = 0$ os autovalores são imaginários puros e o ponto é um de *centro* estável, mas não assintoticamente estável (Fig. 1.1.h).

1.2.3. Estabilidade estrutural e bifurcações

Geralmente, um sistema dinâmico que descreve um sistema físico depende de um ou mais parâmetros de controle. Este tipo de sistema é estruturalmente estável se para qualquer perturbação suficientemente pequena das equações que o descrevem, a solução é topologicamente análoga à solução do sistema sem perturbação (Fiedler-Ferrara e Prado, 1994). Considerando um sistema dinâmico bidimensional temos:

$$\begin{aligned} \dot{x} &= f(x, y, \mu) = f_\mu(x, y) \\ \dot{y} &= g(x, y, \mu) = g_\mu(x, y) \end{aligned} \tag{12}$$

onde μ é o parâmetro de controle. As soluções agora dependem do tempo e do parâmetro μ. Dessa forma, os autovalores e autovetores da matriz Jacobiana também dependem de μ. A variação de μ permite a passagem de um ponto de equilíbrio instável para um estável no valor crítico μ_c que é chamado de **ponto de bifurcação**. Em μ_c o sistema dinâmico perde a estabilidade estrutural e sofre uma bifurcação. O aparecimento de caos neste tipo de sistema está ligado à ocorrência de bifurcações.

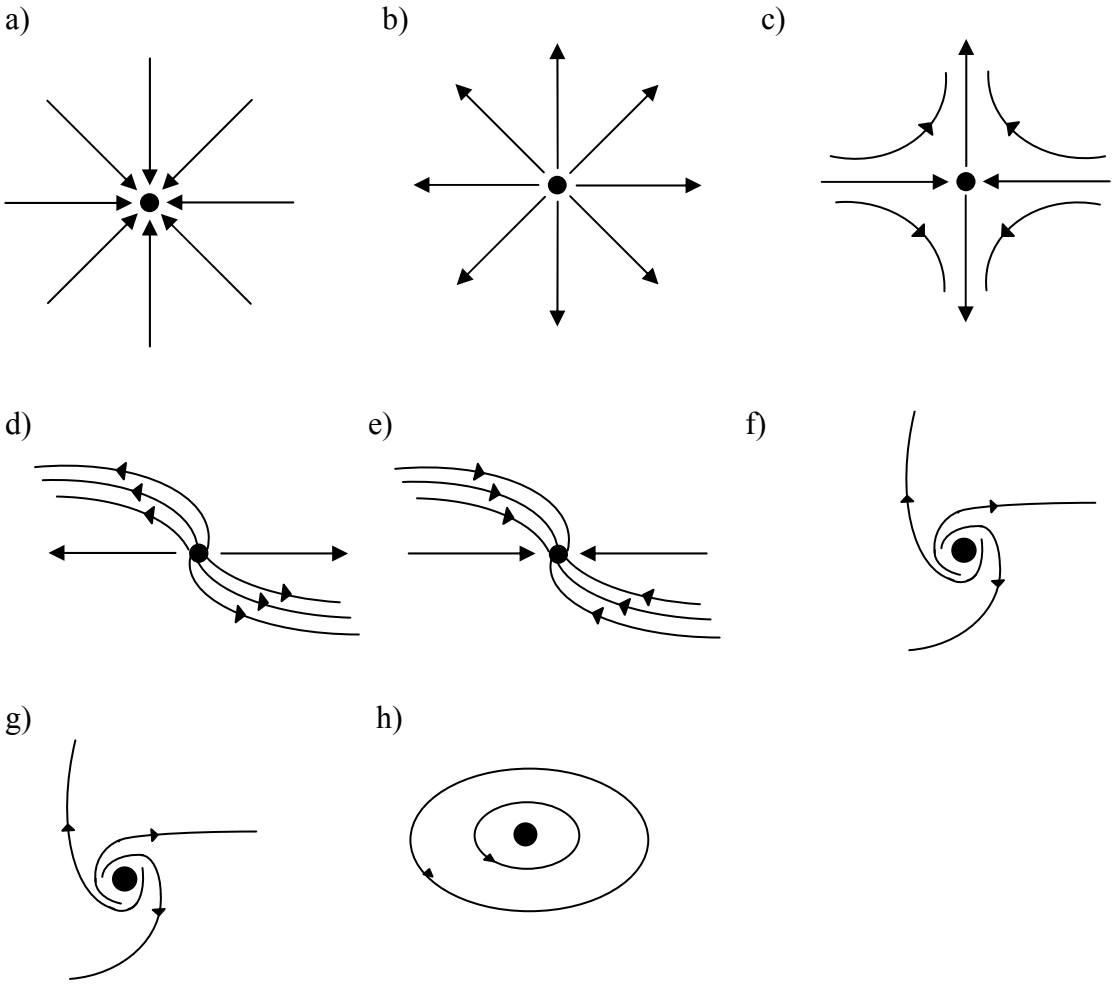


Fig. 1.1. Exemplos típicos da estabilidade dos pontos de equilíbrio em 2 dimensões. a) Nó assintoticamente estável. b) Nó instável. c) Ponto de sela hiperbólico instável. d) Nó impróprio instável. e) Nó impróprio assintoticamente estável. f) Foco instável. g) Foco assintoticamente estável. h) Centro estável, mas não assintoticamente estável.

1.3. Autômatos celulares

A teoria dos autômatos celulares foi introduzida na década de 1950 pelo matemático John von Neumann e Stanislaw Ulam como uma possível idealização para modelar a replicação de organismos biológicos (von Neumann e Burks, 1966). Os modelos baseados em autômatos podem ser aplicados na simulação de sistemas biológicos, sociais e físicos. Os autômatos celulares apresentam três propriedades fundamentais que são paralelismo, localidade e homogeneidade. O paralelismo implica que todos os estados dos sítios são atualizados simultaneamente. O novo estado de um sítio é determinado pelo seu estado anterior e pelo estado de seus vizinhos, ou seja, todas as interações são locais. A evolução temporal é homogênea, pois todos os sítios aplicam as mesmas regras de atualização em relação aos seus vizinhos (Kari, 2005). Os autômatos celulares têm a capacidade de gerar um padrão de comportamento complexo a partir de conjuntos simples de regras. Eles são sistemas com distribuição espacial e temporal discreta, consistindo de um grande número de componentes idênticas, com conectividade local (Wolfram, 1986).

1.3.1. Definição formal

Nesta seção usaremos a notação adotada por Mitchell (1998). O autômato celular consiste de dois componentes que são o espaço celular e a regra de transição. O espaço celular consiste de um reticulado de N sítios idênticos. O conjunto de estados que um sítio pode assumir é dado por Σ e $k = |\Sigma|$ denota o número de estados por sítio. Cada sítio é denotado por um índice i e seu estado em um tempo t é denotado por s_i^t. A vizinhança de um sítio i é determinada pelo estado s_i^t do sítio i e pelos estados dos sítios vizinhos ao sítio i. A regra de transição que é dada por $\phi(\eta_i^t)$ fornece o estado s_i^{t+1} do sítio i, como uma função de η_i^t. A cada passo de tempo, todos os sítios atualizam seus estão de acordo com $\phi(\eta_i^t)$.

1.3.2. Vizinhança

A vizinhança dos autômatos celulares corresponde a uma determinada seleção de sítios próximos. No reticulado unidimensional cada sítio é conectado a r vizinhos de cada lado, onde *r* é um parâmetro definido como raio. A vizinhança de raio 1 consiste nos dois vizinhos mais próximos (Fig. 1.2.a). No reticulado bidimensional, geralmente utiliza-se a vizinhança de Neumann ou a vizinhança de Moore (Wolfram, 1994). A vizinhança de Neumann de raio r para o sítio $\vec{x}_0$ é definida por:

$$N^{v}(\vec{x}_0) = \{\vec{x} : \| \vec{x} - \vec{x}_0 \| \leq r\} \tag{13}$$

onde $\| \cdot \|$ é a norma da soma. Para autômatos celulares bidimensionais temos,

$$N^{v}(x_0, y_0) = \{(x, y) : | x - x_0 | + | y - y_0 | \leq r\} \tag{14}$$

A Fig. 1.2.b mostra a vizinhança do Newman para $r = 1$. A vizinhança de Moore de raio r para o sítio $\vec{x}_0$ é dada por:

$$N^{v}(\vec{x}_0) = \{\vec{x} : | x_i - x_{i0} | \leq r, i = 0,..., d\} \tag{15}$$

onde d é a dimensão do espaço e x_i é a i−ésima componente espacial. A vizinhança de Moore bidimensional é obtida quando $d = 2$. A equação para este tipo de vizinhança é dada por:

$$N^{v}(x_0, y_0) = \{(x, y) : | x - x_0 | \leq r, | y - y_0 | \leq r\} \tag{16}$$

A vizinhança de Moore tridimensional é obtida quando $d = 3$. A equação para este tipo de vizinhança é dada por:

$$N^{v}(x_0, y_0, z_0) = \{(x, y, z) : | x - x_0 | \leq r, | y - y_0 | \leq r, | z - z_0 | \leq r\} \tag{17}$$

As Figs. 1.2.(c) e (d) mostram a vizinhança de Moore com $r = 1$ para $d = 2$ e $d = 3$, respectivamente. A vizinhança também pode ser aleatória (Fig. 1.2.e). Adicionalmente, os autômatos celulares podem ser representados através da teoria dos grafos (O'Sullivan 2001). Neste caso, cada sítio i é associado a um vértice e a aresta representa um relacionamento entre dois sítios i e j. A Fig. 1.2.f mostra a vizinhança dos autômatos celulares baseada em grafos.

1.3.3. Condições de contorno

As condições de contorno mais utilizadas são a periódica e a reflexiva. A condição de contorno periódica é obtida estendendo-se o reticulado periodicamente. A condição de contorno reflexiva é obtida refletindo-se o reticulado em cada limite. Estas duas condições de contorno

podem ser combinadas e pode-se usar em um autômato celular bidimensional, por exemplo, condição de contorno periódica no eixo x e reflexiva no eixo y (Wolfram, 1994). A Fig. 1.3 mostra estes dois tipos de condições de contorno para autômatos unidimensionais.

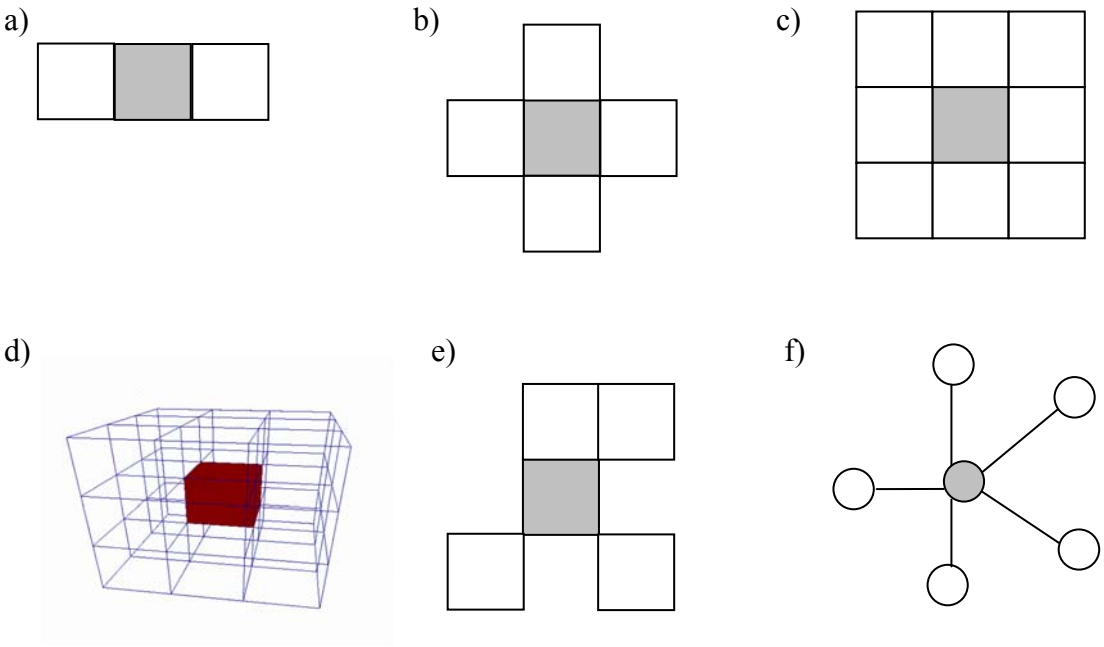


Fig. 1.2. Ilustração de vizinhanças fixas. a) Vizinhança de um reticulado unidimensional de raio 1 – o sítio central possui dois vizinhos. b) Vizinhança de Neumann bidimensional de raio 1 - o sítio central possui quatro vizinhos. c) Vizinhança de Moore bidimensional de raio 1 - o sítio central possui oito vizinhos. d) Vizinhança de Moore tridimensional de raio 1 - o sítio central possui vinte e seis vizinhos. e) Vizinhança aleatória. f) Vizinhança baseada na teoria dos grafos.

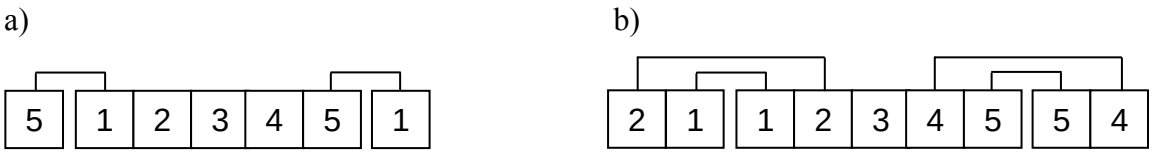


Fig. 1.3. Condições de contorno mais utilizadas para um autômato unidimensional com raio de vizinhança 1. a) Condição de contorno periódica. b) Condição de contorno reflexiva.

1.3.4. Regras de transição

O novo estado de cada sítio depende do seu estado corrente, do estado dos sítios vizinhos e da regra de atualização. Um passo de tempo corresponde à aplicação da regra de transição em todos os sítios do reticulado. Na evolução, cada vez que as regras de transição são aplicadas ao reticulado

completo, uma nova geração é produzida. A evolução temporal dos sítios é feita através de regras locais, usando-se o tipo de vizinhança escolhido. A evolução dos autômatos celulares é conduzida através da aplicação de regras deterministas ou probabilistas. As regras deterministas consideram que o novo estado do sítio depende apenas dos estados dos sítios vizinhos. As regras probabilistas associam para uma dada configuração de sítios vizinhos diferentes saídas e atribuem probabilidades para estas saídas, ou seja, um sítio pode mudar para um novo estado com uma certa probabilidade (Wolfram, 1986).

1.4. Teoria de agentes

Atualmente, a teoria de agentes é considerada uma área de pesquisa muito importante e o principal motivo deste crescente interesse é o seu potencial na modelagem de sistemas complexos. Um agente é considerado uma abstração de alto nível, pois ele é definido em termos de comportamento e não em termos de atributos ou métodos. Este tipo definição o torna mais natural e intuitivo e, assim, amplia o seu uso em diversos tipos de problemas. Dessa maneira, os modelos baseados em agentes são uma extensão dos autômatos celulares devido propriedades mais complexas serem facilmente formuladas e implementadas usando programação orientada a objetos (Segovia-Juarez e cols., 2004).

1.4.1. Definição

O conceito de agente, como uma entidade organizacional, foi introduzido por Carl Hewitt (1977) e ainda não há um consenso sobre uma definição ideal para o agente de software, ou simplesmente agente. As diversas definições estão associadas a diferentes pontos de vista e dependem muito da funcionalidade do agente. Apesar de existirem concordâncias sobre algumas propriedades dos agentes, ainda não existe uma definição totalmente aceita do termo agente. A dificuldade em se encontrar uma única formulação ocorre, pois dentro da comunidade científica, o termo agente é usado em diversos ramos de pesquisa (Nwana, 1996). A seguir serão apresentadas algumas definições de agentes propostas por alguns autores.

Segundo Russell e Norvig (1995), um agente percebe um ambiente através de sensores e atua no mesmo por meio de atuadores. Para Hayes-Roth (1995), um agente pode realizar continuamente três funções, são elas: perceber as condições dinâmicas do ambiente, agir para afetar condições do ambiente e raciocinar para interpretar percepções, resolver problemas, realizar inferências e determinar ações. Os agentes também são definidos como sistemas computacionais que habitam um ambiente dinâmico e/ou complexo, sentem e atuam neste ambiente realizando

tarefas para as quais foram projetados (Maes, 1995). De acordo com Franklin e Graesser (1997) um agente é um sistema que faz parte de um ambiente, podendo sentir e atuar neste ambiente, de acordo com seus próprios propósitos, de modo a alterar o seu futuro. Adicionalmente, para Wooldridge e Jennings (1995) um agente deve implementado utilizando-se conceitos que usualmente são aplicáveis aos seres humanos.

Um agente atua sem intervenção do usuário, tem a capacidade de se comunicar com um humano ou com outro agente de software e de monitorar e perceber o ambiente. Para um agente ser um agente inteligente ele precisa ser capaz de explorar quantidades significativas de conhecimento de domínio, ser tolerante à entrada errada, inesperada ou completamente fora do contexto; ser capaz de usar símbolos e abstrações, possuir comportamento adaptativo e orientado para uma meta, aprender com o ambiente, operar em tempo-real e se comunicar usando linguagem natural (Hayes-Roth, 1995; Wooldridge e Jennings, 1995).

1.4.2. Propriedades de agentes

Alguns aspectos comuns estão presentes nas definições apresentadas pela comunidade científica e as seguintes propriedades podem ser encontradas nos agentes (Franklin e Graesser, 1997; Wooldridge e Jennings, 1995).

- **Autonomia**: os agentes controlam suas ações e seu estado interno;
- **Reatividade**: os agentes percebem seu ambiente e respondem a mudanças que ocorrem nele;
- **Habilidade social**: os agentes interagem com outros agentes e possivelmente com humanos;
- **Pró-atividade**: os agentes não reagem simplesmente em resposta ao ambiente, eles são capazes de exibir comportamentos baseados em objetivos, tomando a iniciativa das suas próprias ações;
- **Mobilidade**: os agentes possuem a habilidade de se mover;
- **Veracidade**: os agentes não comunicam informações falsas de maneira intencional.
- **Benevolência**: os agentes não têm objetivos contraditórios, eles tentam sempre responder ao que lhe é perguntado;
- **Racionalidade**: os agentes sempre agem para alcançar suas metas, e nunca agem contra seus objetivos, pelo menos na medida em que suas crenças o permitam;
- **Flexibilidade**: os agentes não possuem ações rigidamente estabelecidas;

- **Aprendizagem**: os agentes mudam seu comportamento baseando-se em suas experiências prévias.

1.4.3. Arquitetura de agentes

A arquitetura de um agente é determinada pela organização dos agentes dentro de um sistema e como estão estruturados seus relacionamentos e interações Ela pode ser estruturada através de técnicas e algoritmos específicos para a definição de agentes (Knapik, 1998; Maes, 1995; Wooldridge e Jennings, 1995). Os agentes são compostos por módulos que podem interagir através das propriedades que foram implementadas. O conjunto total de módulos e suas interações fornecem informações de como os sensores de dados e o estado interno do agente determinam suas ações e seu futuro estado (Maes, 1995). As arquiteturas podem ser classificadas de acordo com as necessidades da aplicação, dos usuários e do grau de sofisticação ou nível de inteligência dos agentes.

De acordo com Wooldridge e Jennings (1995), as arquiteturas podem ser classificadas em deliberativas, reativas e híbridas. Na arquitetura deliberativa, os agentes possuem um modelo de mundo pré-concebido, porém seu estado pode ser alterado através da comunicação com outros agentes da rede e através da percepção do ambiente. Neste modelo, o agente estima que ações sejam necessárias para alcançar um determinado objetivo, e então executa as ações que o levarão ao seu objetivo. Na arquitetura reativa, os agentes não possuem inteligência, raciocínio simbólico complexo e representação interna do mundo, eles possuem apenas mapeamentos de situações e respostas associadas. Este modelo possui um forte acoplamento entre percepção e ação. As decisões tomadas por este tipo de agente são baseadas em uma quantidade limitada de informação. Assim, quando ocorre uma mudança no ambiente, o agente executa a ação correspondente. A arquitetura híbrida é uma mistura das arquiteturas deliberativa e reativa com o objetivo de tornar os agentes mais funcionais.

Segundo Knapik e Johnson (1998), a complexidade da arquitetura de um sistema pode ser classificada em arquitetura simples, moderada e complexa. A arquitetura simples é composta por somente um único tipo de agente. A arquitetura moderada é composta por alguns agentes que realizam as mesmas tarefas, mas possuem diferentes usuários e podem residir em máquinas diferentes. A arquitetura complexa é composta por diferentes tipos de agentes, cada um com certa autonomia, podendo cooperar e estar em diferentes plataformas.

1.4.4. Sistemas Multiagentes

Os sistemas multiagentes estudam o comportamento de um grupo de agentes autônomos que cooperam para resolver problemas que vão além das capacidades de solução de cada agente (Wooldridge, 2002). Eles são considerados dinâmicos, pois o ambiente pode ser alterado pelos agentes (Russell e Norvig, 1995). Neste tipo de sistema, a organização é essencial, pois coloca restrições aos comportamentos dos agentes, procurando estabelecer um comportamento de grupo coeso. Em um cenário multiagente típico, existe interação entre os agentes. Esta interação pode ser vista como um estímulo ambiental ou como uma comunicação inter-agentes (Jennings e Wooldridge, 1998; Wooldridge, 2002).

Nos sistemas multiagentes, os agentes são concebidos independentemente de um problema particular, as interações entre os agentes são definidas de modo a serem usadas em situações genéricas, não existe um controle centralizado para a resolução de um problema e a decomposição de tarefas para solucionar um dado problema pode ser feita pelos próprios agentes. Dessa forma, a utilização dos sistemas multiagentes em modelagem possui algumas vantagens, são elas: viabilizar sistemas adaptativos e evolutivos, pois este tipo de sistema possui a capacidade de se adaptar a novas situações; tirar proveito de ambientes heterogêneos e distribuídos, devido a agentes com arquiteturas diferentes e que funcionam em plataformas diferentes, poderem cooperar na resolução de problemas; e modelar sistemas complexos, pois decomposição de problemas e a atribuição dos sub-problemas a agentes permitem um alto nível de abstração e independência entre as partes do sistema (Jennings e Wooldridge, 1998; Wooldridge, 2002).

1.4.5. Modelos de multiagentes

Um modelo computacional baseado em agentes consiste na criação de uma população de agentes autônomos, capazes de reproduzir as características do componente que estiver sendo simulado. Neste tipo de modelagem, os agentes devem possuir estado interno, perceber e agir no ambiente. No estado interno estão armazenadas as propriedades dos agentes que se relacionam com seus objetivos, motivações situação atual. Para que os agentes possam trafegar de modo autônomo é necessário que eles sejam capazes de perceber o ambiente no qual estejam inseridos. Adicionalmente, o comportamento do agente define como ele deve agir e/ou realizar determinadas tarefas. Para se modelar o comportamento necessita-se da identificação e análise das ações que serão reproduzidas pelos agentes. O ambiente da simulação corresponde ao espaço disponível para movimentação e interação dos agentes (Russell e Nörvig, 1995; Wooldridge, 2002). Este ambiente pode ser representado por uma rede regular ou complexa.

1.5. Redes complexas

A teoria das redes complexas engloba conceitos provenientes da teoria dos grafos, mecânica estatística, e sistemas complexos. As redes constituem estruturas onde processos dinâmicos naturais e artificiais podem ser simulados. Ela é utilizada em diversas áreas da ciência e esse caráter multidisciplinar permite esta teoria cobrir aplicações em física, química, matemática, biologia, psicologia, sociologia, linguística, dentre outras. A computação é responsável pelas ferramentas utilizadas na modelagem, simulação e tratamento das bases de dados (Barabási, 2003).

1.5.1. Conceitos básicos

Matematicamente, uma rede é representada por um grafo. Um grafo pode ser definido por um par $G = \{V, L\}$, onde V é o conjunto de vértices (nós) e L é o conjunto de arestas (conexões, ligações) que conecta dois vértices de G. Uma rede é dirigida ou orientada quando as conexões entre os pares de vértices são orientadas e é não-dirigida ou não-orientada quando tais conexões não são orientadas. O termo aresta é denominado para as conexões das redes não-orientadas e o termo arco para as conexões de redes orientadas (Albert e Barabási, 2002; Costa e cols., 2007; Dorogovtsev e Mendes, 2002). Se as arestas possuem intensidade, um peso é associado a cada aresta. Então, a rede passa a ser representada por um trio $G = \{V, L, P\}$ onde P é o conjunto de pesos (Costa e cols., 2007). Dois vértices são denominados adjacentes quando existe uma aresta que os une. Em uma rede regular todos os vértices possuem o mesmo número de arestas. Por outro lado, em uma rede complexa os vértices possuem um número diferente de arestas (Albert e Barabási, 2002; Costa e cols., 2007).

As redes podem ser representadas através de diagramas, nos quais os vértices são representados por pontos e as arestas por linhas que ligam esses pontos (Boccaletti e cols., 2006). Em termos computacionais, as redes podem ser armazenadas através de matriz ou lista de adjacência (Boccaletti e cols., 2006), por exemplo. A matriz adjacente (Netto, 2006) é representada por uma matriz $|B| x |B|$, tal que o elemento da matriz $d_{ij} = 1$ se existir uma aresta entre b_i e b_j, e 0 (zero) se não existir uma ligação entre b_i e b_j. Em uma lista de adjacência (Netto, 2006), apenas os pares de vértices (i, j) que possuem arestas são armazenados. Assim, o uso da lista de adjacência permite maior economia de memória em relação ao uso de matriz de adjacência quando as redes são esparsas. A lista de adjacência possui um elemento relacionado à intensidade das ligações (i, j, w_{ij}) quando as conexões possuem peso. Adicionalmente, pode-se utilizar uma matriz de pesos ao invés da matriz de adjacência para armazenar os pesos das conexões entre os vértices (Costa e cols.,

2007). A Fig. 1.4 mostra o mapeamento de uma rede não-orientada e de uma orientada em lista e matriz de adjacência.

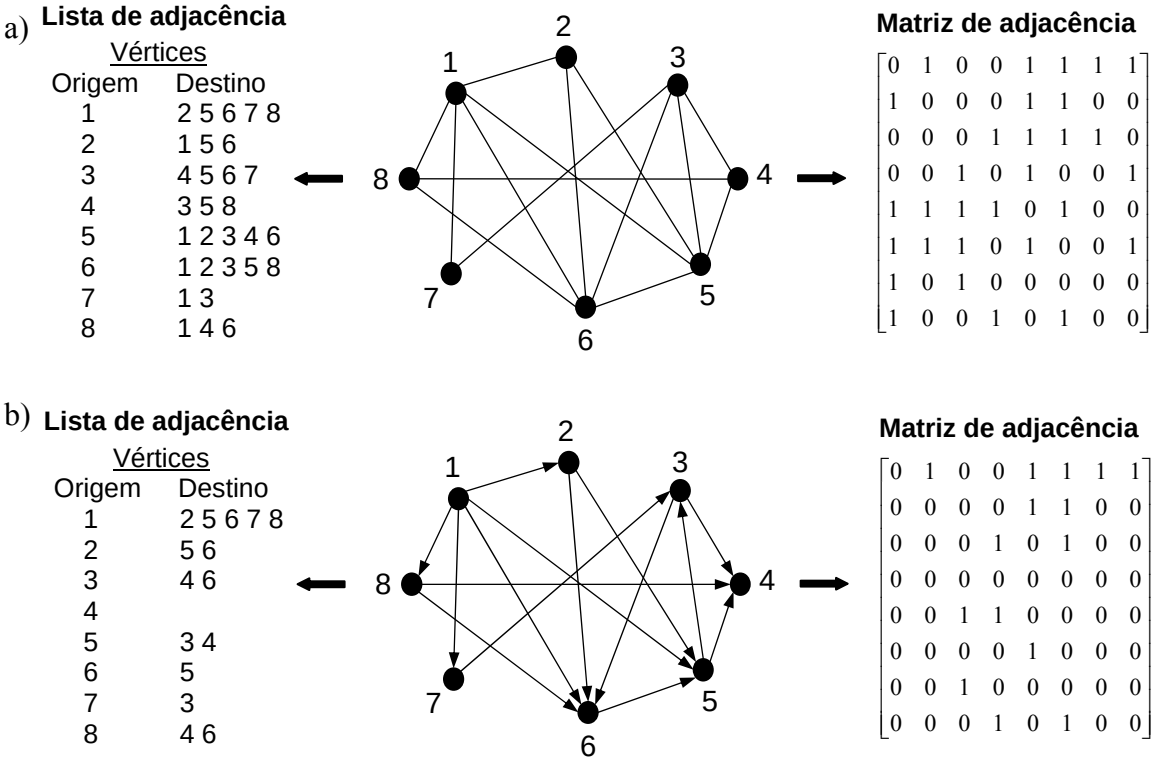


Fig. 1.4. Ilustração de um grafo e a sua representação por lista de adjacência e matriz de adjacência. a) Grafo não-orientado. b) Grafo orientado.

1.5.2. Índices das redes complexas

A taxonomia das redes complexas pode ser feita agrupando-se em uma mesma categoria as redes cujas medidas estruturais forneçam valores estatisticamente semelhantes. A escolha dos índices para se caracterizar a rede é de extrema importância, pois um conjunto reduzido de índices ou a consideração de índices que forneçam resultados semelhantes, pode acarretar classificações equivocadas. Deste modo, o conhecimento e a utilidade dos índices é fundamental para caracterização, comparação, classificação e modelagem de redes complexas (Costa e cols., 2007). A seguir serão descritos os índices mais comuns na análise de uma rede.

1.5.2.1. Índices relacionados à conectividade

O **grau** k_i de um vértice é geralmente definido como o número de conexões que o vértice possui. Os vértices muito conectados, ou seja, concentradores, são essenciais na formação da

estrutura das redes complexas, já que a sua retirada pode causar a fragmentação da rede, resultando em componentes não-conectados. Esse fenômeno está relacionado às falhas aleatórias e ataques direcionados a redes. O **grau médio** da rede $< k >$ é calculado pela média aritmética dos graus de cada vértice. A **distribuição de graus** em uma rede é caracterizada por uma função de distribuição $P(k)$, onde esta função descreve a probabilidade de um determinado vértice selecionado aleatoriamente tenha k arestas. A distribuição de graus na rede é obtida através do histograma (Albert e Barabási, 2002; Boccaletti e cols., 2006; Costa e cols., 2007; Dorogovtsev e Mendes, 2002).

1.5.2.2. Índices relacionados à distância

O comprimento do caminho que conecta dois vértices i e j é dado pelo número de arestas ao longo deste caminho. O **caminho mínimo** entre dois vértices i e j, é dado pelo número de arestas do menor caminho que os conecta. O **caminho mínimo médio** é a média da distância entre dois vértices medida sobre todos os pares de vértices. O **diâmetro** de uma rede é definido como a máximo caminho mínimo entre quaisquer dois pontos da rede (Albert e Barabási, 2002; Boccaletti e cols., 2006; Dorogovtsev e Mendes, 2002). Alguns vértices ou arestas estão situados entre muitos caminhos mínimos e isso significa que há um grande fluxo que passa por este vértice ou aresta. O **grau de intermediação** de um vértice (Costa e cols., 2007) é usado para medir o tráfego que passa em um dado vértice, ou seja, este índice mede o quanto um vértice está no caminho entre outros vértices. Similarmente, o **grau de intermediação** de uma aresta (Girvan e Newman, 2002) mede quanto uma aresta está no caminho entre pares de vértices. Ao se remover um vértice ou aresta com alto grau de intermediação pode ocorrer ruptura na estrutura da rede, surgindo componentes não-conectados, que são formados por vértices altamente conectados entre si. A expressão que define o grau de intermediação de um vértice é dada por:

$$B(v) = \sum_{i,j} \frac{\sigma_{ij}(v)}{\sigma_{ij}} \tag{18}$$

onde, $\sigma_{ij}(v)$ é o número de caminhos mínimos entre os vértices i e j que passam pelo vértice ou aresta v e σ_{ij} é o número total de caminhos mínimos entre i e j. A média do grau de intermediação pode ser utilizada como um índice para medir a caracterização global da rede,

$$\langle B \rangle = \frac{1}{N} \sum_{i} B_i \tag{19}$$

onde N é o número total de vértices ou arestas.

1.5.2.3. Índice relacionado a ciclo

O **coeficiente de aglomeração** de um vértice é dado pela razão entre o número de arestas entre os vizinhos de um dado vértice i e o número máximo de arestas entre esses vizinhos para eles estarem completamente conectados entre si (Albert e Barabási, 2002; Boccaletti e cols., 2006; Costa e cols., 2007; Dorogovtsev e Mendes, 2002). Matematicamente, o coeficiente de aglomeração pode ser dado por:

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \tag{20}$$

Onde, E_i é dado pelo número de arestas dos vértices adjacentes ao vértice i e k_i é o número de arestas do vértice i. A expressão $k_i(k_i - 1)/2$ representa o número máximo de arestas que os vizinhos de um vértice, com k_i arestas, pode ter. O coeficiente de aglomeração médio é dado pela razão entre a soma dos coeficientes de aglomeração de cada vértice e número total de vértices na rede. Então,

$$C = \frac{1}{N}\sum_{1}^{N} C_i \tag{21}$$

Se o coeficiente de aglomeração for zero significa que os nós adjacentes a i não estão conectados entre si, mas se for um significa que todos os vértices adjacentes a i estão conectados entre si.

1.5.2.4. Dimensão fractal

A investigação da auto-similaridade das redes complexas é realizada através da renormalização usando o método de contagem de caixas, como mostrado na Fig. 1.5. Para uma dada rede G com N vértices e N_B caixas, a caixa é o conjunto de vértices onde todas as distâncias entre quaisquer dois vértices é menor que ℓ_B. Cada caixa é substituída por um vértice e este procedimento é repetido até que a rede inteira seja reduzida a um único vértice. Para ℓ_B = 1, N_B é igual ao tamanho da rede G. Quando N_B = 1, temos que $\ell_B \geq \ell_B^{\max}$, onde $\ell_B^{\max}$ é o diâmetro da rede adicionando-se um. As redes complexas que possuem topologia fractal podem ser caracterizadas pelas seguintes relações de escala:

$$\frac{N_B(\ell_B)}{N} \sim \ell_B^{-d_B} \qquad \text{e} \qquad \frac{k_B(\ell_B)}{k_{hub}} \sim \ell_B^{-d_k} \tag{22}$$

onde $k_B(\ell_B)$ é o grau de cada caixa e k_{hub} é o grau do vértice mais conectado. A componente d_B é a dimensão fractal e d_k é o grau dos expoentes das caixas (Song e cols, 2005; 2006; 2007).

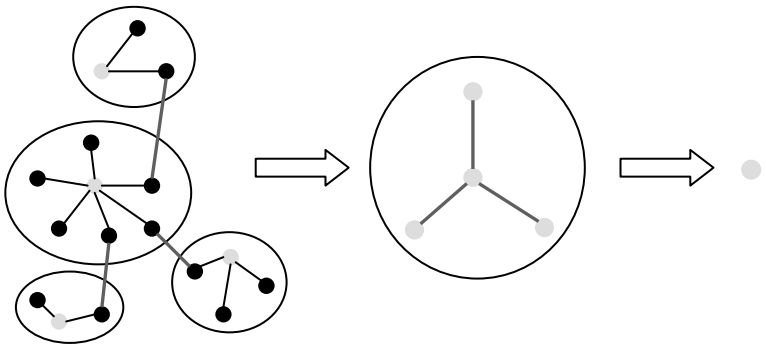


Fig. 1.5. Auto-similaridade das redes complexas. Esta rede possui 17 vértices que são renormalizados com $N_B(\ell_B) = 4$ caixas de tamanho $\ell_B = 3$. Ela possui topologia fractal, pois as caixas são conectadas através dos vértices não-concentradores.

1.5.2.5. Identificação de comunidades

Muitas redes reais possuem comunidades, ou seja, as conexões são densamente distribuídas entre vértices que pertencem a um mesmo grupo e esparsamente distribuídas entre os vértices de grupos diferentes (Fig. 6). As comunidades são formadas por vértices que possuem alguma similaridade entre si e a presença de comunidades na rede corresponde à natureza hierárquica de sistemas complexos (Costa e cols., 2007). Um algoritmo bastante popular usada na identificação de comunidades é baseado no grau de intermediação das arestas da rede (Girvan e Newman, 2002). Este algoritmo remove as arestas responsáveis pelas conexões entre as comunidades. Inicialmente, é calculado o grau de intermediação de todas as arestas na rede. Em seguida, a aresta com maior grau de intermediação é removida e o grau de intermediação para todas as arestas afetadas pela remoção é recalculado. Este processo é repetido até que a rede esteja completamente dividida em vértices isolados. As comunidades da rede são identificadas ao se remover as arestas com maior grau de intermediação, pois grupos isolados aparecem.

A qualidade de uma divisão particular da rede é medida através da **modularidade**, Q, proposta por Newman e Girvan (Newman e Girvan, 2004). A divisão de uma rede em k comunidades é obtida considerando-se uma matriz simétrica e, cujos componentes $e_{i,j}$ fornecem as frações de arestas que conectam vértices presentes na comunidade i aos vértices presentes na

comunidade j. O traço desta matriz $Tre = \sum_i e_{i,i}$ representa a fração de arestas que conectam vértices dentro da mesma comunidade. O traço por si só não é um bom indicador da qualidade da divisão da rede, pois se todos os vértices estiverem presentes em uma única comunidade o valor do traço será máximo ($Tre = 1$) e isto não fornece informações sobre as comunidades na rede. Para resolver esta limitação, utiliza-se o somatório dos elementos das linhas (ou colunas) da matriz $a_i = \sum_j e_{ij}$, que representa a fração de arestas que conectam dois vértices presentes em uma comunidade i. Portanto, a **modularidade** Q pode ser definida da seguinte forma,

$$Q = \sum_i (e_{ii} - a_i^2) = Tre - \left\| e^2 \right\| \tag{23}$$

onde, $\left\| e \right\|$ é a soma de todos elementos da matriz e.

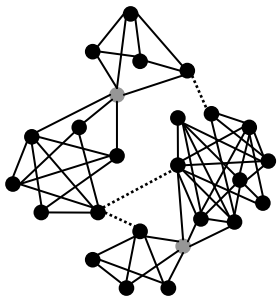


Fig. 1.6. Ilustração de uma rede com quatro comunidades. Os vértices cinzas possuem alto grau de intermediação, pois através deles duas comunidades se sobrepõem. O fluxo entre as comunidades ocorre pelas arestas pontilhadas e estas arestas também possuem alto grau de intermediação.

1.5.3. Modelos de redes complexas

No século vinte a teoria dos grafos se desenvolve, tornando-se mais algorítmica e, assim, é possível aplicá-la ao estudo de redes complexas, surgindo, assim, as redes aleatórias. Esta teoria foi proposta por Paul Erdos e Alfréd Rényi (Erdös e Rényi, 1959). O modelo de Erdös e Rényi (ER) apresenta um grafo aleatório como um conjunto de N vértices e n arestas, sendo que cada par de vértices é conectado com uma probabilidade p (Fig. 1.7.a). O número de arestas é dado por $n = pN(N-1)/2$ e as arestas são distribuídas aleatoriamente na rede. Este tipo de arquitetura possui uma distribuição de graus de Poisson e vértices desconectados podem ser encontrados neste tipo de rede.

Em 1998, Watts e Strogatz (WS) (Watts e Strogatz, 1998) propuseram uma rede que se situa entre uma rede regular e uma rede aleatória. A rede de mundo pequeno inicia-se com uma rede regular formada por N vértices conectados aos seus primeiros k vizinhos. Cada vértice é aleatoriamente reconectado com uma probabilidade p; a transição entre uma rede completamente regular ($p = 0$) e uma rede aleatória ($p = 1$) pode ser obtida pela variação de p. Uma variante do modelo de WS foi proposto por Newman e Watts (NW) (Newman e Watts, 1999a 1999b), onde somente arestas são adicionadas entre os vértices (Fig. 1.7b). O modelo de NW é mais fácil de ser analisado, pois ele não possui vértices desconectados, devido o padrão da rede regular ser mantido.

Barabási e Albert (BA) propuseram uma rede em que a distribuição de graus seguia uma lei de potência, ou seja, poucos vértices possuem muitas conexões e muitos vértices possuem poucas conexões. O modelo de BA começa com um número pequeno de vértices (m_0) e a cada passo de tempo novos vértices (m) são adicionados e se conectam aos vértices existentes por uma ligação preferencial (Fig. 1.7c). A probabilidade (Π) de um vértice se conectar depende do grau (k_i) do vértice (i) (Albert e Barabási, 2002), tal que:

$$\Pi(k_i) = \frac{k_i}{\sum_i k_i} \tag{24}$$

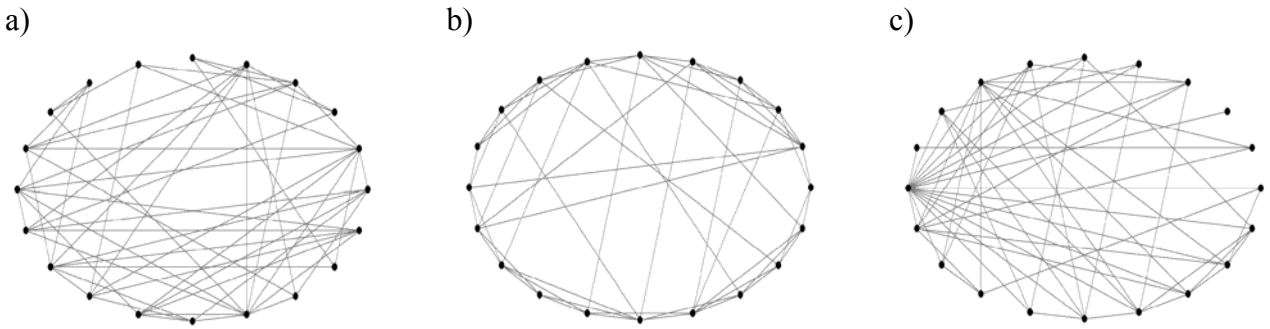


Fig. 1.7. Estrutura de redes complexas gerada pelo Pajek (Batagelj and Mrvar, 1998) com 20 vértices e 56 arestas. a) Rede aleatória. b) Rede de mundo pequeno. c) Rede livre de escala.

1.6. Conclusões

Neste trabalho discutimos as técnicas mais utilizadas na modelagem de sistemas complexos. Os modelos baseados em autômatos celulares ou agentes possuem distribuição espacial e temporal discreta diferentemente do baseado em equações diferenciais que possui estas distribuições contínuas. A teoria de agentes é uma extensão dos autômatos celulares, pois propriedades mais

complexas são formuladas e implementadas mais facilmente usando programação orientada a objetos. O ambiente em que ocorre a simulação pode ser representado por uma rede regular ou complexa para modelos baseados em autômatos ou agentes. Em uma rede regular todos os vértices possuem o mesmo número de arestas. Numa rede complexa o número de arestas para cada vértice é diferente e esta distribuição é o que caracteriza os distintos tipos de redes.

Agradecimentos

Este trabalho foi financiado pela FAPESB.

Referências

Adamatzky, A. (1994) Identification of Cellular Automata. Taylor and Francis, London.

Albert, R., Barabási, A.-L. (2002) Statistical mechanics of complex networks. Rev. of Mod. Phys. 74, 47-97.

Amaral, L.A.N., Ottino, J.M. (2004) Complex networks. Eur. Phys. J. B, 38, 147–162.

Andrade, R.F.S., Pinho, S.T.R., Fraga, S.C., Tanajura, A.P.M. (2002) New toppling dynamics for the self-organized critical rainfall model. Physica A, 314, 405 – 410.

Barabási, A.-L. (2003) Linked: how everything is connected to everything else and what it means for bBusiness, science, and everyday life. Plume, New York, NY.

Bar-Yam, Y. (1997) Dynamics of Complex Systems. Perseus Books, Reading, MA.

Batagelj, V., Mrvar, A. (1998) Pajek-program for large network analysis. Connections, 21, 47–57.

Braun, M. (1992) Differential equations and their applications: an introduction to applied mathematics. Fourth ed., Springer, USA.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.-U. (2006) Complex networks: structure and dynamics. Phys. Rep., 424, 175 – 308.

Costa, L.D., Rodrigues, F.A., Travieso, G., Villas Boas, P.R.V. (2007) Characterization of complex networks: A survey of measurements. Adv. Phys., 56, 167 – 242.

Dorogovtsev, S.N., Mendes, J.F.F. (2002) Evolution of networks. Adv. Phys., 51, 1079-1187.

Erdös, P., Rényi, A. (1959) On random graphs. Publ. Math., 6, 290-297.

Fiedler-Ferrare, N., do Prado, C.P.C. (1994) Caos: uma introdução. Edgard Blucher, São Paulo, Brasil.

Franklin, S., Graesser, A. (1997) Is it an agent, or just a program?: A taxonomy for autonomous agents. Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, 1193, 21-35, Springer, Berlin.

Galvão, V., Miranda, J.G.V. (2008) A computational model for cancer growth by using complex networks. Physica A, 387, 5279–5286.

Galvão, V., Miranda, J.G.V., Ribeiro-dos-Santos, R. (2008) Development of a two-dimensional agent-based model for chronic chagasic cardiomyopathy after stem cell transplantation. Bioinformatics, 24, 2051-2056.

Galvão, V., Miranda, J.G.V. (2009) Modeling the Chagas' disease after stem cell transplantation. Physica A, 388, 1747-1754.

Girvan, M., Newman, M.E.J. (2002) Community structure in social and biological networks. PNAS, 99, 7821–7826.

Hayes-Roth, B. (1995) An Architecture for Adaptive Intelligent Systems. Artif. Intell., 72, 329-365.

Hewitt, C. (1977) Viewing control structures as patterns of passing messages. Artif. Intell., 8, 323-364.

Ioannou, P.A, Pitsillides, A. (2007) Modeling and control of complex systems, CRC Press.

Jennings, N. R., Wooldridge, M. J. (1998) Agent Technology: foundations, applications, and markets. Springer, Berlin.

Kari, J (2005) Theory of cellular automata: a survey. Theor. Comput. Sci., 334, 3 – 33.

Knapik, M., Johnson, J (1998) Developing intelligent agents for distributed systems: exploring architecture, technologies, and applications, McGraw-Hill, New York, NY.

Maes, Pattie (1995) Artificial life meets entertainment: life like autonomous agents. Commun. ACM, 38, 108-114.

Mitchell, M. (1998). Computation in cellular automata: a selected review. In Gramss, T., Bornholdt, S., Gross, M., Mitchell, M., Pellizzari, T., Nonstandard Computation, Weinheim: VCH Verlagsgesellschaft, 95-140.

Mitchell, M. (2006) Complex systems: Network thinking. Artif. Intell., 170, 1194–1212.

Muñoz, P.C., da Silva, M.F., Miranda, J.G.V., Claro, F., Diniz, R.G. (2007) Influence of government controls over the currency exchange rate in the evolution of Hurst's exponent: An autonomous agent-based model. Physica A, 386, 786-790.

Murray, J.D. (2002) Mathematical biology: I. an introduction. Third ed., Springer, New York.

Netto, P.O.B. (2006) Grafos: teoria, modelos, algoritmos. Blucher, Quarta ed., São Paulo, 2006.

Newman, M.E.J., Girvan, M. (2004) Finding and evaluating community structure in networks. Phys. Rev. E, 69, 026113.

Newman, M.E.J., Watts, D.J. (1999a) Renormalization group analysis of the small-world network model. Phys. Let. A. 263, 341–346.

Newman, M.E.J., Watts, D.J. (1999b) Scaling and percolation in the small-world network model. Phys. Rev. E, 60, 7332-7342.

Nwana, H. S. (1996) Software agents: an overview. Know. Eng. Rev., 11, 205-244.

Oliveira, F.S.B.F. (2002) Wave climate modelling south of Rio de Janeiro in Brazil. Cont. Shelf Res., 22, 2021–2034.

O'Sullivan, D. (2001) Graph-cellular automata: a generalised discrete urban and regional model. Environ. Plan B, 28, 687-705.

Ottino, J.M. (2004) Engineering complex systems. Nature, 427, 399.

Pinho, S.T.R., Andrade, R.F.S. (1998) An Abelian model for rainfall. Physica A, 255, 483- 495.

Queirós, S.M.D., Curado, E.M.F., Nobre, F.D. (2007) A multi-interacting-agent model for financial markets. Physica A, 374, 715-729.

Reis, E.A., Santos, L.B.L., Pinho, S. T. R. (2009) A cellular automata model for avascular solid tumor growth under the effect of therapy. Physica. A, 388, 1303-1314.

Ribeiro, L.C. et al. (2007) Dynamic stability in random and scale-free B-lymphocyte networks. Phys. Rev. E, 75, 031911.

Russell, S., Nörvig, P. (1995) Artificial intelligence: a modern approach – The Intelligent Agent Book. Prentice Hall, New Jersey, USA.

Segovia-Juarez, J.L., Ganguli, S., Kirschner, D. (2004). Identifying control mechanisms of granuloma formation during M. tuberculosis infection using an agent-based model. J. Theor. Biol., 231, 357-376.

Song, C., Havlin, S., Makse, H.A. (2005) Self-similarity of complex networks. Nature, 433, 392-395.

Song, C., Havlin, S., Makse, H.A. (2006) Origins of fractality in the growth of complex networks. Nature Phys., 275-281.

Song, C., Gallos, L.K., Havlin, S., Makse, H.A. (2007) How to calculate the fractal dimension of a complex network: the box covering algorithm. J. Stat. Mech., P03006.

Thompson, J.M.T., Stewart, H.B. (2002) Nonlinear Dynamics and Chaos. Second ed., John Wiley and Sons, England.

Verhulst, F. (2000) Nonlinear differential equations and dynamical systems. Springer, second ed., Berlin- Heidelberg, New York.

von Neumann, J., Burks, A.W. (1966) Theory of self-reproducing automata. University of Illinois Press, Champaign, IL, USA.

Watts, D.J., Strogatz, S.H. (1998) Collective dynamics of small-world network. Nature. 393, 440-442.

Whitesides, G.M., Ismagilov, R.F. (1999) Complexity in Chemistry. Science, 284, 89-92.

Wolfram, S. (1986) Theory and Applications of Cellular Automata. World Scientific, Singapore.

Wolfram, S. (1994) Cellular automata and complexity: collected papers. Addison-Wesley.

Wooldridge, M.J. (2002) An introduction to multiagent systems. John Wiley and Sons, New York, USA.

Wooldridge, M.J., Jennings, N.R. (1995) Intelligent agents: theory and practice. Knowl. Eng. Rev., 10, 115-152.

Zorzenon dos Santos, R.M., Coutinho, S.G. (2001) The Dynamics of the Evolution HIV Infection: A Cellular Automata Approach. Phys. Rev. Lett., 87, 16810.

Zorzenon dos Santos, R.M., Pinho, S.R.T., Pio, C.F., da Silva, P.C.A. (2007) On the study of the dynamical aspects of parasitemia in the blood cycle of Malaria. Eur. Phys. J. Special Topics, 1, 125-134.

CHAPTER 2

Development of a two-dimensional agent-based model for chronic chagasic cardiomyopathy after stem cell transplantation

Abstract

Motivation: A significant issue in stem cell therapy is to understand the role of this type of cell in the tissue regeneration. To explain this mechanism, an experimental study has quantified that the bone marrow cell transplantation decreases the number of inflammatory cells and reduces the fibrosis area in chagasic mice. Using this experimental data, we have developed an agent-based computational model to investigate the regeneration of the chronic chagasic cardiomyopathy after bone marrow stem cell transplantation.

Results: Our model includes six different types of agents: inflammatory cell, fibrosis area, cardiomyocyte, proinflammatory cytokinetumor necrosis factor-α, *Trypanosoma cruzi* parasite and bone marrow stem cell. This latter promotes apoptosis in inflammatory cells, reduction in the fibrosis area and can differentiate into cardiomyocyte. Proinflammatory cytokine tumor necrosis factor-α can increase the fibrosis area and *T. cruzi* can increase the number of inflammatory cells. Our results for both apoptosis of inflammatory cells and reduction in the fibrosis area were compared with experimental data. They suggest that the concentration pattern is the most important factor to characterize the kinetics of cardiac tissue regeneration after bone marrow stem cell transplantation.

Availability: The source code of our software is available online at www.vivas.ufba.br/bone/bone.zip

2.1. Introduction

Chronic chagasic cardiomyopathy is an endemic illness affecting millions of people in the world. The etiological agent of this disease is the hemoflagellate parasite *Trypanosoma cruzi*. This protozoan is transmitted by more than 100 species of insects of the family *Ruduviidae*, subfamily

Triatominae or by blood transfusion. According to the World Health Organization, there are 16–18 millions of people infected only in South America. This disease is still incurable, however most of the contaminated people remain asymptomatic, i.e. they have an indeterminate form and around 30% of them develop a chronic inflammatory disease, with cardiac and/or digestive complications. The clinical symptoms of the heart pathology are cardiomyopathy, heart failure and arrhythmia (Andrade and Andrews, 2005; Coura et al., 2002; Higuchi et al., 2003; Kelly, 2000; Zacks et al., 2005).

Chronic chagasic cardiomyopathy is characterized by a diffuse inflammatory reaction composed mainly of mononuclear cells, and a severe fibrosis (Higuchi et al., 2003; Soares et al., 2001). The number of *T. cruzi* parasites in the chronic phase of this disease is disproportionately low in relation to the intensity of inflammation. However, recent studies have verified high frequencies of *T. cruzi* antigens in this disease, providing that the occurrence of this parasite is associated with myocardial inflammation (Higuchi et al., 2003). The development of fibrosis is caused by an elevate production in the number of proinflammatory cytokines tumor necrosis factor-α (TNF-α) in the heart (Soares et al., 2004).

The contribution of the scientific literature in computational models for parasites infection includes interaction between *T. cruzi* parasites and antibodies during the acute phase of Chagas infection (Condat et al., 2003; Isasi et al., 2001; Sibona and Condat, 2002), competitive parasite–antibody interaction in the intracellular andextracellular phase of the Chagas' disease (Sibona et al., 2005), cell-mediated immune response to Chagas' disease (Nelson and Velasco-Hernández, 2002) and dynamics of Plasmodium falciparum invasion of the human erythrocyte cells (Ferrer et al., 2007; Hoschen et al., 2000; McKenzie and Bossert, 2005; Peleg et al., 2002).

Also, a great deal of research has been carried out to investigate the kinetics of stem cell transplantation (Bianco et al., 2001; Grove et al., 2004; Sell, 2004; Temple, 2001). A stem cell is a singular type of cell that can renew itself and possesses ability to divide in multiple types of specialized cells often for indefinite periods (Sell, 2004). The discovery of the plasticity of adult bone marrow cells has allowed new perspectives for the treatment of incurable chronic disease. Adult bone marrow cells include two types of bone marrow stem cells, i.e. hematopoietic stem cells and mesenchymal stem cells. Hematopoietic stem cells give rise to all blood lineages and mesenchymal stem cells give rise to bone, cartilage and fat. Experimental evidences have shown that adult bone marrow stem cell can differentiate into other types of cells, including skeletal muscle, cardiac muscle, hepatocytes, keratinocytes and neurons (Bianco et al., 2001; Grove et al., 2004; Sell, 2004).

Diverse experimental models have confirmed the regenerative potential of bone marrow stem cells. The transplant of bone marrow stem cells improves the injured tissue because this type

of stem cell can promote apoptosis in non-normal cells. Also, bone marrow stem cells can differentiate into a normal cell of the tissue (Sell, 2004). An experimental model to describe the effects of adult bone marrow cell transplantation in dilated cardiomyopathy was developed by Soares et al. (2004). In this model, bone marrow cells containing bone marrow stem cells were injected into chronic chagasic mice leading to a considerable reduction in the inflammatory infiltrate and in the fibrosis area. The authors suggest that bone marrow stem cells promote apoptosis of inflammatory cells and reduction in the fibrosis area.

Recently, different theoretical models have been proposed to understand the stem cell differentiation and proliferation. The models developed for describing these processes are based on differential equations (Lemon et al., 2007; Tannenbaum et al., 2005), co-clustering latent variable models (Joung et al., 2006) and Bayesian network (Woolf et al., 2005). However, we can also represent stem cells by using autonomous agents to evaluate the influence of this population. For instance, the different types of cells in the chronic chagasic cardiomyopathy can be represented by different types of agents. Thus, we propose in this article a two-dimensional agent-based model for chronic chagasic cardiomyopathy regeneration after bone marrow stem cell transplantation.

2.2. Computational model

A two-dimensional lattice is employed to represent the cardiac tissue and the state of the sites is chosen and distributed on a regular lattice. Each site represents the region of the space in which only a type of autonomous agent can occupy in the chagasic tissue. The time evolution is equally run in the entire lattice and each site changes its state according to a local rule, which depends only on the adjacent neighbors. In addition, our model has different types of agents to describe the chagasic cardiac tissue. In this model, each type of agent represents a different type of cell. An agent lives in an environment in which it can identify the types of agents through their attributes and can interact with other agents (Wooldridge and Jennings, 1995). There are also empty spaces to allow the movement of some types of cells. We have developed a software program to simulate the chronic chagasic cardiomyopathy regeneration after bone marrow stem cell transplantation. The source code was developed in C++ language and it is available at http://www.vivas.ufba.br/bone/bone.zip.

2.2.1. State of the sites

The computational model generates a two-dimensional rectangular lattice consisting of a grid with 500×500 points in size, which represents a block of virtual chagasic cardiac tissue. This

lattice possesses 250 000 sites and periodic boundary conditions. Initially, the parameters of this model are the total number of agents, initial fraction of fibrosis area, initial fraction of inflammatory cells, initial fraction of bone marrow stem cells, initial fraction of TNF-α and initial fraction of *T. cruzi* parasites. The fraction of cardiomyocytes is given by the difference between the total number of agents and the total number of the others types of cells. Here, we use the experimental data obtained by Soares et al. (2004) to simulate the initial fraction of inflammatory cells and initial fraction of fibrosis area. Thus, in our simulations we employ the same quantity for these two parameters.

The quantity of each type of agent is calculated by multiplying the initial fraction of each type of agent by the total number of agents. The quantity of bone marrow stem cells, *T. cruzi* parasites and TNF-α is constant during the simulation. The initial distribution of the agents is heterogeneous. Initially, some fibrosis areas are randomly placed and the others are placed in empty spaces in the neighborhood of a fibrosis area with a certain probability. This probability depends on the quantity of fibrosis areas around the empty space. The same occurs to the inflammatory cells. The probabilities are given by:

$$P_\beta = \frac{T_\beta}{8} \qquad (2.1)$$

$$P_\gamma = \frac{T_\gamma}{8} \qquad (2.2)$$

where, P_β is the probability of an empty space being occupied by a fibrosis area, T_β is the total number of fibrosis areas surrounding the empty space, P_γ is the probability of an empty space being occupied by an inflammatory cell and T_γ is the total number of inflammatory cells surrounding the empty space.

Cardiomyocytes, bone marrow stem cells, TNF-α, and *T. cruzi* have a random distribution. The distribution pattern of all types of agents is based on experimental data of Soares et al. (2004). In this case, the fibrosis area and inflammatory cells have a non-uniform density; therefore, we use a similar concentration distribution in these types of agents.

2.2.2. Neighborhood

Inflammatory cells, cardiomyocytes, TNF-α, *T. cruzi* parasites and bone marrow stem cells can move through the empty spaces. Each type of cell (agent) is chosen randomly making it to jump

to empty space. These types of agents move randomly using the Moore neighborhood. In this type of neighborhood, each cell has eight neighbors situated in positions north, south, east, west, northeast, southeast, northwest and southwest (Wolfram, 1986). Only the fibrosis area is fixed.

2.2.3. Transition rules of states

The time evolution depends on a set of rules that determines the state of all sites. The state of the sites is determined through the simulated stem cell properties. This computational model simulates the properties of apoptosis and differentiation. Bone marrow stem cells can promote apoptosis of inflammatory cells, reduction of the fibrosis area (Soares et al., 2004) and can differentiate into cardiomyocytes (Orlic et al., 2001). Also, TNF-α can increase the fibrosis area (Soares et al., 2004) and *T. cruzi* can increase the number of inflammatory cells (Higuchi et al., 2003). Then, the transition rules of this computational model are as follow:

• In the neighborhood of an inflammatory cell, if the number of bone marrow stem cells is different from zero, the apoptosis of this inflammatory cell occurs (Soares et al., 2004), i.e., the site of the inflammatory cell changes into empty space.

• In the neighborhood of a fibrosis area, if the number of bone marrow stem cells is different from zero, the reduction in this fibrosis area occurs (Soares et al., 2004), i.e., the site of the fibrosis area changes into empty space.

• In the neighborhood of an empty space site, if there is a cardiomyocyte and the number of bone marrow stem cells is greater than a determined value, the differentiation of bone marrow stem cells into cardiomyocyte occurs (Orlic et al., 2001), i.e., the empty space changes into cardiomyocyte.

• In the neighborhood of an empty space site, if the number of inflammatory cells and *T. cruzi* parasites is different from zero, the empty space site changes into an inflammatory cell (Higuchi et al., 2003).

• In the neighborhood of an empty space site, if the number of TNF-α and fibrosis areas is different from zero, the empty space site changes into a fibrosis area (Soares et al., 2004).

2.3. Results and discussion

A total of 20 simulations were performed to estimate the best parameters that describe the kinetics of bone marrow stem cells in the chagasic cardiac tissue. All the parameters were varied and the parameter optimization was done using the best fit with the experimental data. All the steps of the model were equally followed in all simulations. The data of fibrosis area and inflammatory cells were normalized to enable a comparison between computational and experimental results. The normalization of experimental and computational results is done in the same way. All the data are divided by the initial value of the respective dataset. Then, the normalized data correspond to the fraction of cells in relation to the first month. This can be calculated by using the equation:

$$N'_i = \frac{N_i}{N_o} \qquad (2.3)$$

where N'_i represents the normalized data, N_i the original data and N_o the data in the first month. As defined in the model rules, bone marrow stem cells can promote apoptosis of inflammatory cells, reduction in the fibrosis area and generation of cardiomyocytes. TNF-α increases the fibrosis area and *T. cruzi* parasites increase the number of inflammatory cells.

To compare the predictions of the agent-based model, we carried out a fit of the experimental data (Soares et al., 2004). In this experiment, 8-week-old female or male BALB/c mice were used for *T. cruzi* infection. These mice were infected by the inoculation of *T. cruzi* trypomastigotes and treated 6 month later with adult bone marrow cells obtained from normal BALB/c mice by the intravenous route. These bone marrow cells contained several different types of cells, including bone marrow stem cells. Bone marrow cells were used to treat chagasic mice. These mice were sacrificed 1, 2, 3, 5 and 6 months after treatment for measurement of inflammatory cells per square millimeter and percentage of fibrosis area in the heart. In this experiment, the authors measured the fibrosis area and counted the number of inflammatory cells only 1 month after bone marrow cell transplantation.

Several studies have emphasized that the quantity of parasite in the chronic phase of Chagas' disease is disproportionately low (Higuchi et al., 2003). However, in the experiment of Soares et al. (2004) parasites could not be observed either in the blood or in heart sections 1, 2 and 4 weeks and 2 to 6 months after bone marrow cell transplantation, by using optical microscopy. Also, in this experiment the migration of transplanted bone marrow cells was demonstrated, but the number of bone marrow stem cell was not quantified. Also, the number of TNF-α in the chagasic cardiac tissue was not quantified.

Figure 2.1a shows heart sections of chronic chagasic mice treated with saline and bone marrow cells obtained by Soares et al. (2004). In these heart sections, the regions of fibrosis and inflammatory cells are greater and more concentrated in the non-transplanted mice. Figure 2.1b–d shows a graphic visualization of our computational model. In this representation, for clarity, the states of the sites have different patterns. The distribution of the different types of agents in our model describes the concentration pattern of fibrosis area and inflammatory cells in the chagasic cardiac tissue. The model evolution simulates the dissolution and reduction of fibrosis area and inflammatory cells. Therefore, initially the regions of fibrosis and inflammatory cells are more concentrated, but they do not possess a uniform density. The concentration pattern of the different types of agents has a significant importance because it is responsible for the chagasic tissue regeneration.

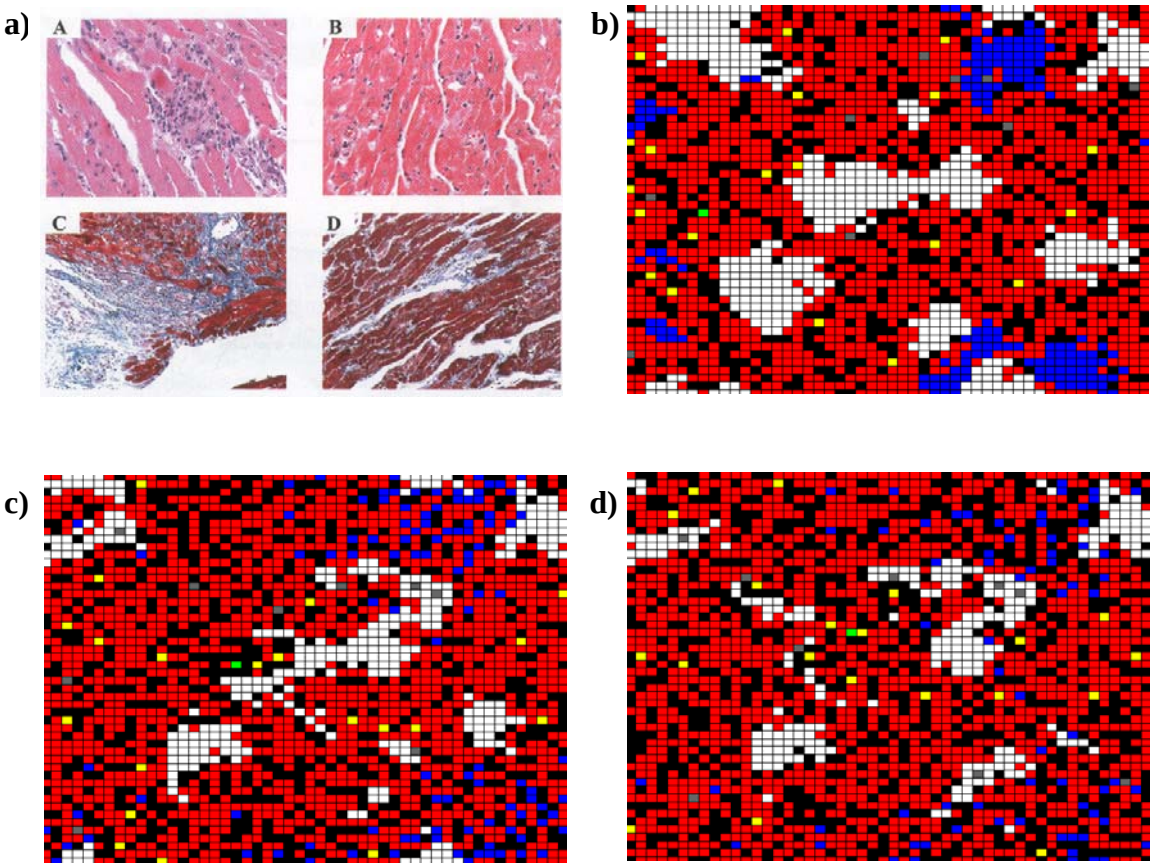


Fig. 2.1. Comparison of the chagasic cardiac tissue section from Soares *et al.* (2004) with our computational model at different time steps. a) Heart sections of the chronic chagasic BALB/c mice treated with saline (control) and bone marrow cells. Two month later control mice (A and C) or bone marrow cells treated mice (B and D) were sacrificed. In these sections, fibrosis area appears in white, inflammatory cells in blue and cardiomyocytes in red. In these heart sections, the regions of fibrosis area and inflammatory cells are greater and more concentrate in the control mice. (Reprinted from Am. J. Pathol. 2004 164: 441-447 with permission from the American Society for Investigative Pathology). b) Graphic representation of our computational model at different time steps. This two-dimensional rectangular lattice consists, for clarity, of 50 x 50 points. The initial fraction of the lattice-sites occupied (So) is equal to 0.8, the initial fraction of inflammatory cells is equal to 0.08, the initial fraction of fibrosis area is equal to 0.22, the initial fraction of *T. cruzi* (Tc) is equal to 0.00055, the initial fraction of TNF-α is equal to 0.0045, the initial fraction of bone marrow stem cells (Bmst) is equal to 0.01, and the number of bone marrow stem cells required to occur differentiation of one bone marrow stem cell into a cardiomyocyte (Dcm) is greater than 2. Empty spaces are represented by black patterns, fibrosis areas by white, cardiomyocytes by red, bone marrow stem cells by yellow, TNF-α by grey and *T. cruzi* by green. Time step $t = 0$. c) Time step $t = 50$. d) Time step $t = 100$.

In this computational model, 500 time steps correspond to the experimental data of 1 month; therefore, each time step corresponds to~1.44 h of 'real time'. Hence, we have performed 2500 time steps. In all the simulations, we use the initial fraction of inflammatory cells equaling 0.08, and the initial fraction of fibrosis area equaling 0.22. We have used these values because these are the fractions obtained by Soares et al. (2004), 1 month after bone marrow cell transplantation. We have compared our computational data with the experimental data 1, 2, 3, 5 and 6 months after treatment.

In Figure 2.2, we show the kinetics of apoptosis of inflammatory cells and reduction in the fibrosis area for different fractions of lattice-site occupation. Our results indicate that the smaller the number of empty spaces, the smaller the final fraction of inflammatory cells and fibrosis area. This happens because it is more difficult for a *T. cruzi* parasite to meet an inflammatory cell when there are few empty spaces. Therefore, when the number of empty spaces is small, the increase in the number of inflammatory cells occurs more slowly. This is related to the random movement of inflammatory cells and *T. cruzi*. The same occurs for TNF-α.

In Figure 2.2a, the experimental curve of inflammatory cells Soares et al. (2004) shows an increase that probably occurs since the mice are infected with *T .cruzi* and, consequently, continue generating inflammatory cells. In Figure 2.2b, the experimental curve of fibrosis area (Soares et al., 2004) first decreases, then increases and in the sixth month decreases again. The experimental curve of fibrosis area increases in the third month because there is TNF-α in the cardiac tissue of mice; therefore, the development of fibrosis area continues. However, the reduction of fibrosis in the sixth month can perhaps be attributed to the effect of decreases in the number of inflammatory cells, TNF-α-producing cells in the heart.

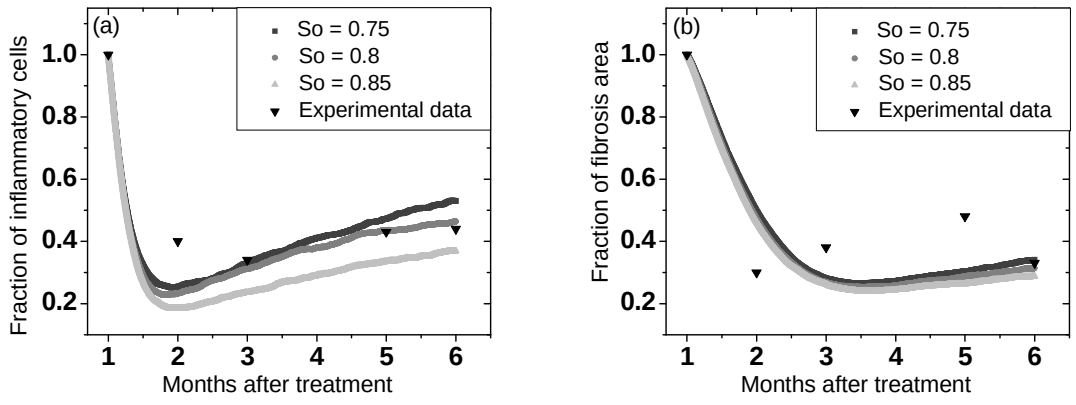


Fig. 2.2. Kinetics of chagasic tissue regeneration for different fractions of lattice-site occupation. The other parameters used were Tc = 0.00055, TNF-α = 0.0045, Bmst = 0.01, Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Apoptosis of inflammatory cells. (b) Reduction in the fibrosis area. (See Fig. 2.1 for the meaning of the labels).

The variation of the initial fraction of bone marrow stem cells for apoptosis of inflammatory cells and reduction in the fibrosis area is shown in the Figure 2.3. The fit for experimental data are given in Figure 2.3a and b. In addition, the data of our computational model also increase because it possesses *T. cruzi* and TNF-α. The presence of *T. cruzi* and TNF-α is essential for the simulation of chagasic cardiac regeneration. This fact is evidenced because in the absence of *T. cruzi*, the number of inflammatory cells tends to zero and in the absence of TNF-α, the fibrosis area tends to zero. This vanishing is due to the presence of the bone marrow stem cell.

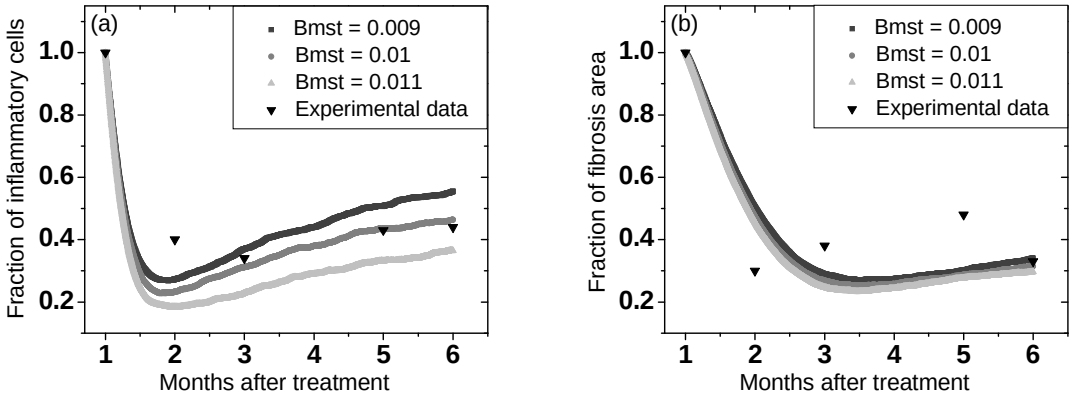


Fig. 2.3. Kinetics of chagasic tissue regeneration for different fractions of bone marrow stem cells. The other parameters used were So = 0.8, Tc = 0.00055, TNF-α = 0.0045 and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Apoptosis of inflammatory cells. (b) Reduction in the fibrosis area. (See Fig. 2.1 for the meaning of the labels).

In Figure 2.4a, we vary the initial fraction of *T. cruzi* and in Figure 2.4b the initial fraction of TNF-α. We can observe that a little difference in the initial fraction of the *T. cruzi* changes the behavior of the apoptosis of inflammatory cells. The behavior of the reduction in the fibrosis area is less affected by the variation of the fraction of TNF-α. This happens because it is more difficult for a bone marrow stem cell to penetrate in a place with high density than with low density.

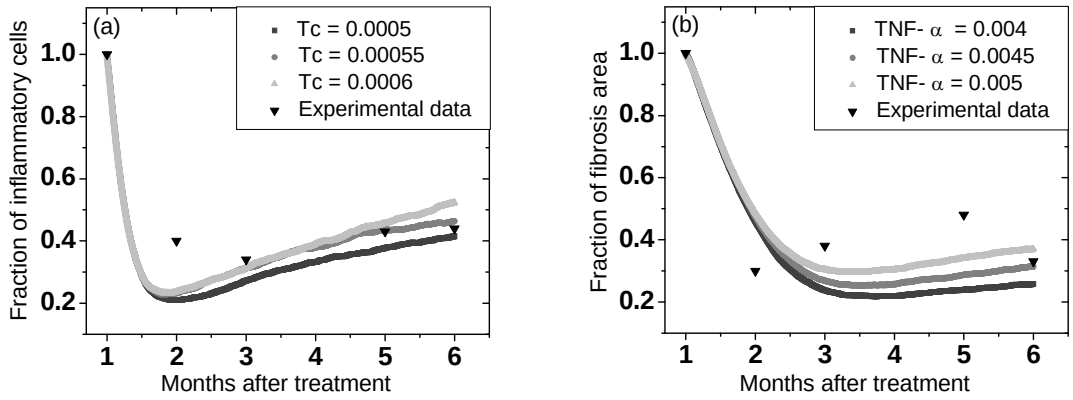


Fig. 2.4. Kinetics of chagasic tissue regeneration for different fractions of *T. cruzi* and TNF-α. The other parameters used were So = 0.8, Bmst = 0.01 and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Apoptosis of inflammatory cells. In this simulation we use TNF-α = 0.0045. (b) Reduction in the fibrosis area. In this simulation we use Tc = 0.00055. (See Fig. 2.1 for the meaning of the labels).

Figure 2.5 shows the differentiation of the bone marrow stem cells into cardiomyocyte. In all the simulations, the increase is linear. However, in the neighborhood of an empty space site, if there is a cardiomyocyte and the number of bone marrow stem cells is greater than 2, the total number of cardiomyocytes grows around 4.5% and if the number of bone marrow stem cells is greater than 3, the total number of cardiomyocytes grows around 0.6%. It is well known that bone marrow stem cells can differentiate into a normal cell of the tissue (Sell, 2004). As the experimental data of Soares et al. (2004) do not quantify the number of new cardiomyocytes, we can only suggest that the increase of cardiomyocytes is less realistic when the number of bone marrow stem cells is greater than 3 because it is very small.

The best parameters of our computational model to describe the chagasic tissue regeneration were chosen using the best fit with the experimental data. Thus, the parameters selected are the fraction of the number of sites occupied equaling 0.8, fraction of *T. cruzi* parasites equaling 0.00055, fraction of TNF-α equaling 0.0045, fraction of bone marrow stem cells equaling 0.01 and the number of bone marrow stem cells required for the differentiation of a bone marrow stem cell into a cardiomyocyte to occur is greater than 2.

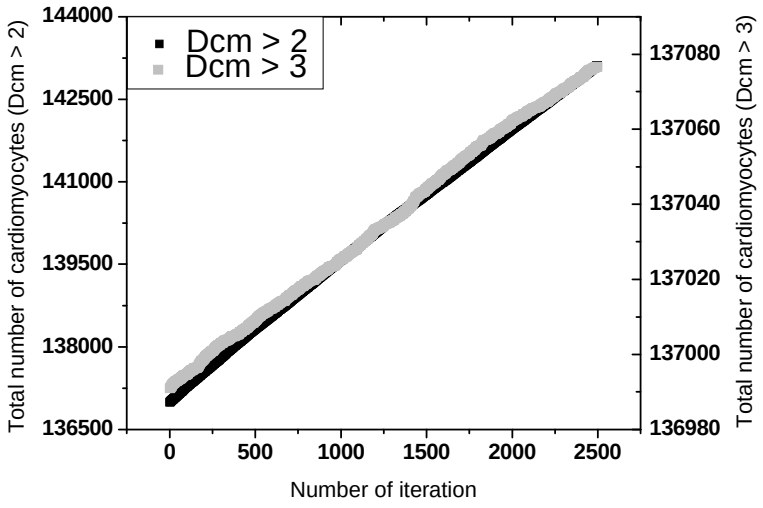


Fig. 2.5. Kinetics of chagasic tissue regeneration for different number of bone marrow stem cells required for differentiation of a bone marrow stem cell into a cardiomyocyte. (See Fig. 2.1 for the meaning of the labels).

2.4. Conclusion

In this article, we have presented a two-dimensional agent-based model to investigate the regeneration of the chronic chagasic cardiomyopathy after bone marrow stem cell transplantation. Our computational model can simulate some stem cells properties including apoptosis and differentiation. Apoptosis occurs in inflammatory cells and the bone marrow stem cells can

differentiate in cardiomyocytes. Also, there is reduction in the fibrosis area. The fundamental hypothesis is that the kinetics of chagasic tissue regeneration has a concentration pattern and this can be modeled and represented by using an appropriate computational model. The main aim of our model is to understand the participation of different types of cells in the cardiac tissue regeneration.

Our results were compared with experimental data, excepting for the differentiation of bone marrow stem cells in cardiomyocytes. The results also suggest that the concentration pattern of fibrosis area and inflammatory cells is the most important factor in the kinetics of chronic chagasic cardiomyopathy regeneration after bone marrow stem cell transplantation. In addition, the initial fraction of bone marrow stem cells and *T. cruzi* parasites affects the rate of apoptosis of inflammatory cells. Finally, the initial fraction of bone marrow stem cells and TNF-α affects the rate of reduction in the fibrosis area.

Acknowledgements

This work has been supported by the Brazilian agency FAPESB and the authors thank Roberto Rivelino (Instituto de Física, Universidade Federal da Bahia, Brazil) for a careful review of this article.

Conflict of Interest: none declared.

References

Andrade, L.O., Andrews, N.W. (2005) The *Trypanosoma cruzi*–host-cell interplay: location, invasion, retention. Nat. Rev. Microb., 3, 819–823.

Bianco, P. et al. (2001) Bone marrow stromal stem cells: nature, biology, and potential applications. Stem Cells, 19, 180–192.

Condat, C.A. et al. (2003) Parasite-antibody competition in chagas disease. Comments Theor. Biol., 8, 587–607.

Coura, J.R. et al. (2002) Emerging chagas disease in Amazonian Brazil. Trends Parasitol., 18, 171–176.

Ferrer, J. et al. (2007) Individual-based model and simulation of Plasmodium falciparum infected erythrocyte in vitro cultures. J. Theor. Biol., 248, 448–459.

Grove, J.E., Bruscia, E., Krause, D.S. (2004) Plasticity of bone marrow-derived stem cells. Stem Cells, 22, 487–500.

Higuchi, M.L. et al. (2003) Pathophysiology of the heart in Chagas' disease: current status and new developments. Cardiovasc. Res., 60, 96–107.

Hoschen, M.B. et al. (2000) Mathematical modeling of the within-host dynamics of Plasmodium falciparum. Parasitology, 121, 227–235.

Isasi, S.C., Sibona, G.J., Condat, C.A. (2001) A simple model for the interaction between T. cruzi and its antibodies during Chagas infection. J. Theor. Biol., 208, 1–13.

Kelly, J.M. (2000) A B-cell activator in chagas disease. Nat. Med., 6, 865–866.

Joung, J.-G. et al. (2006) Identification of regulatory modules by co-clustering latent variable models: stem cell differentiation. Bioinformatics, 22, 2005–2011.

Lemon, G. et al. (2007) Mathematical modelling of human mesenchymal stem cell proliferation and differentiation inside artificial porous scaffolds. J. Theor. Biol., 249, 543–553.

McKenzie, F.E., Bossert, W.H. (2005) An integrated model of *Plasmodium falciparum* dynamics. J. Theor. Biol., 232, 411–426.

Nelson, P., Velasco-Hernández, J.X. (2002) Modeling the immune response to parasitic infections: Leishmaniasis and Chagas disease. Comments Theor. Biol., 6, 161.

Orlic, D. et al. (2001) Bone marrow cells regenerate infarcted myocardium. Nature, 410, 701–705.

Peleg, M. et al. (2002) Modelling biological processes using workflow and Petri Net models. Bioinformatics, 18, 825–837.

Sell, S. (2004) Stem Cell Handbook. Humana Press, Totowa, NJ.

Sibona, G.J., Condat, C.A. (2002) Dynamic analysis of a parasite population model. Phys. Rev. E, 65, 031918.

Sibona, G.J., Condat, C.A., Isasi, S.C. (2005) Dynamics of the antibody-*T. cruzi* competition during Chagas infection: prognostic relevance of intracellular replication. Phys. Rev. E, 71, 020901.

Soares, M.B.P. et al. (2001) Modulation of chagasic cardiomyopathy by interleukin-4: dissociation between inflammation and tissue parasitism. Am. J. Pathol., 159, 703–709.

Soares, M.P.B. et al. (2004) Transplanted bone marrow cells repair heart tissue and reduce myocarditis in chronic chagasic mice. Am. J. Pathol., 164, 441–447.

Tannenbaum, E. et al. (2005) Evolutionary dynamics of adult stem cells: comparison of random and immortal-strand segregation mechanisms. Phys. Rev. E, 71, 041914.

Temple, S. (2001) The development of neural stem cells. Nature, 414, 112–117.

Zacks, M.A. et al. (2005) An overview of chagasic cardiomyopathy: pathogenic importance of oxidative stress. An. Acad. Bras. Cienc., 77, 695–715.

Wolfram, S. (1986) Theory and Applications of Cellular Automata. World Scientific, Singapore.

Wooldridge, M., Jennings, N.R. (1995) Intelligent agents: theory and practice. Knowl. Eng. Rev., 10, 115–152.

Woolf, P.J. et al. (2005) Bayesian analysis of signaling networks governing embryonic stem cell fate decisions. Bioinformatics, 21, 741–753.

CHAPTER 3

Modeling the Chagas' disease after stem cell transplantation

Abstract

A recent model for Chagas' disease after stem cell transplantation is extended for a three-dimensional multi-agent-based model. The computational model includes six different types of autonomous agents: inflammatory cell, fibrosis, cardiomyocyte, proinflammatory cytokine tumor necrosis factor-α, *Trypanosoma cruzi*, and bone marrow stem cell. Only fibrosis is fixed and the other types of agents can move randomly through the empty spaces using the three-dimensional Moore neighborhood. Bone marrow stem cells can promote apoptosis in inflammatory cells, fibrosis regression and can differentiate in cardiomyocyte. *T. cruzi* can increase the number of inflammatory cells. Inflammatory cells and tumor necrosis factor-α can increase the quantity of fibrosis. Our results were compared with experimental data giving a fairly fit and they suggest that the inflammatory cells are important for the development of fibrosis.

3.1. Introduction

The discovery of the pluripotency of adult bone marrow stem cells has opened new perspectives for the treatment of patients with chronic chagasic cardiomyopathy. This disease is caused by the hemoflagellate parasite *Trypanosoma cruzi* and one of the leading causes of heart failure in Latin America. Chagas' disease is transmitted by an insect of the subfamily *Triatominae*, or by blood transfusion. Most of the *T. cruzi*-infected individuals remain asymptomatic. This subclinical state is called the indeterminate form of Chagas' disease. However, around 30% of the *T. cruzi*-infected individuals develop a cardiac complication, in a late phase of the disease. Chronic chagasic cardiomyopathy is characterized by a diffuse inflammatory reaction and a severe fibrosis (Andrade and Andrews, 2005; Higuchi et al., 2003; Soares et al., 2001; Soares et al., 2004). In the chronic phase of the disease, the presence of *T. cruzi* is associated with chronic inflammatory

response (Higuchi et al., 2003). The elevated production of the proinflammatory cytokine tumor necrosis factor-α (TNF-α) causes the fibrosis growth (Soares et al., 2004).

The regenerative potential of bone marrow stem cell transplantation has been under investigation in experimental models of ischemic cardiomyopathy and chronic chagasic cardiomyopathy (Soares et al., 2004; Sell, 2004). A stem cell is a particular type of cell that can renew itself and possesses ability to divide in many different types of specialized cells (Sell, 2004; Grove et al., 2004; Bianco et al., 2001). Experimental evidences have shown that the transplant of bone marrow stem cell can promote apoptosis in non-normal cells and can differentiate into a normal cell of the tissue (Sell, 2004). An experimental model to describe the effects of adult bone marrow transplant in the chronic chagasic cardiomyopathy was developed by Soares et al. (2004). In this model, bone marrow cells were injected into chronic chagasic mice leading to a considerable reduction in the number of inflammatory cells and in the fibrosis area. The authors suggest that bone marrow stem cells can promote apoptosis of inflammatory cells and fibrosis regression.

The scientific contribution in computational models for parasites includes interaction between *T. cruzi* and antibodies during the acute phase of Chagas infection (Sibona and Condat, 2002; Isasi et al., 2001), competitive parasite-antibody interaction in the intracellular and extracellular phase of the Chagas' disease (Sibona et al., 2005), cell-mediated immune response to Leishmaniasis and Chagas' disease (Nelson and Velasco-Hernández, 2002), dynamics of *Plasmodium falciparum* invasion of the human erythrocyte cells (Ferrer et al., 2007), mathematical model of immune response in cutaneous Leishmaniasis (de Almeida and Moreira, 2007), cellular automata model for the dynamical aspects of parasitemia in the blood cycle of malaria (Zorzenon dos Santos et al., 2007) and time delay effect in the *Plasmodium* of *Physarum polycephalum* by using the living coupled oscillator system (Takamatsu et al., 2000).

In recent years, some different theoretical models have been proposed to understand the kinetics of stem cells. These models are based on cellular automata (Agur et al., 2002), agent systems (d'Inverno and Saunders, 2005), differential equations (Tannenbaum et al., 2005), co-clustering latent variable models (Joung et al., 2006), Bayesian network (Nagarajan et al., 2005; Woolf et al., 2005), lattice (Zhdanov, 2008), stochastic lattice (Sahimi et al., 1997), and other mathematical models (Adimy et al., 2008a; Adimy et al., 2008b; Pisu et al., 2007; Pisu et al., 2008). In this way, we have recently developed a two-dimensional agent-based model for Chagas' disease (Galvão et al., 2008) after bone marrow stem cell transplantation. The main aim of this computational model was to understand the participation of different types of cells in the chagasic tissue regeneration after bone marrow stem cell transplantation. The two-dimensional model only simulated the heart sections and the fibrosis regression had small discrepancies between the computational and experimental data.

In order to better understand the behavior of fibrosis after bone marrow stem cell transplantation, the model was extended here for a three-dimensional multi-agent-based model. Also, the transition rule for fibrogenesis was changed to include the influence of inflammatory cells in the development of fibrosis. The strong point of this type of study lies in the possibility of validating hypothesis about the behavior of different types of cells. In our computational model the hypothesis are presented by rules that simulate the cellular kinetics. The reasons for developing this model are the increase of realism due to the reproduction of chagasic cardiac tissue in three-dimensions and to better comprehend the increase of fibrosis. The advantages of this model in relation to our previous model (Galvão et al., 2008) are to simulate the volume of the heart and the inclusion of inflammatory cells in the fibrosis progression.

3.2. Computational model

The description of our computational model follows the standard ODD protocol (Overview, Design concepts, and Details) for individual-based and agent-based models (Grim et al., 2006).

3.2.1. Purpose

The purpose of this computational model is to comprehend how the spatial change affects the contribution of different types of cells in the chronic chagasic cardiomyopathy regeneration after bone marrow stem cell transplantation.

3.2.2. State variables and scales

The three-dimensional rectangular lattice is composed by a grid of cubic cells, called sites. The lattice consists of 100 x 100 x 100 points in size, which represents the chagasic cardiac tissue. This rectangular lattice possesses periodic boundary conditions and each site represents the region of space which only a type of autonomous agent can occupy. To make the model more realistic each agent represents a single cell. The system description is discrete in time and space and all actions occur at constant intervals called time steps. Each autonomous agent can be present or absent in a particular site. In a site, if the number of agents is zero; the site is referred to as empty space. Therefore, the empty space site is equivalent to space in real system that is not occupied by an agent. In addition, each type of autonomous agent represents a different type of cell.

In our model six different types of agents are distinguished: inflammatory cell, fibrosis, cardiomyocyte, TNF-α, *T. cruzi*, and bone marrow stem cell. The parameters of this model are the

total number of autonomous agents and the initial fractions of fibrosis, inflammatory cells, bone marrow stem cells, TNF-α, and *T. cruzi*. The fraction of cardiomyocytes is given by the difference between the total number of agents and the total number of the other types of cells. Our model represents time as discrete time steps with a fixed duration and the time evolution is equally run in the entire lattice. In our model, 1000 time steps correspond to the experimental data (Soares et al., 2004) of five months; therefore one time step corresponds to approximately 3.6 h of ``real time".

3.2.3. Process overview and scheduling

At each time-step, the same sequence of actions occurs sequentially and builds the configuration of the next step scenario. The inflammatory cells, cardiomyocytes, TNF-α, *T. cruzi*, and bone marrow stem cells can jump to empty space. These types of agents move randomly using the three-dimensional Moore neighborhood. In this case, each site has twenty six adjacent neighbors (Wolfram, 1986). Only the fibrosis is fixed. The procedures applied by the different types of agents are described by five rules in the section ``Submodels".

3.2.4. Design concepts

Emergence: Population dynamics and structure emerge from the behavior of cells. The chagasic tissue regeneration is represented entirely by local and deterministic rules, describing, for example, the apoptosis of inflammatory cells and fibrosis reduction.
Sensing: All agents are assumed to know all their characteristics and they apply the rules according to them. Agents sense the state of the other agents and act according to the state of their first neighbors.
Interaction: The system dynamics is driven by direct interactions between inflammatory cell, fibrosis, cardiomyocyte, TNF-α, *T. cruzi*, and bone marrow stem cell. The interactions are local, i.e., in the neighborhood of an agent.
Observation: At each time step, significant variables such as output and information about the system state are statistically analyzed. These variables include: number of inflammatory cells, quantity of fibrosis and number of cardiomyocytes.
Stochasticity: The initialization is the only part that the stochasticity is considered.

3.2.5. Initialization

We use the experimental data obtained by Soares et al. (2004) to determine the initial fraction of inflammatory cells and initial fraction of fibrosis. In this experiment, fibrosis and inflammatory cells have a non-uniform density and form clusters. Therefore, the initial concentration distribution of the autonomous agents should be heterogeneous. Thus, to simulate this type of density, initially some fibrosis are randomly placed and after the other have a probability of being placed in the neighborhood of a fibrosis already established. This probability depends on the quantity of fibrosis around the empty space. The same occurs to the inflammatory cells. The probabilities are given by:

$$P_{\beta} = \frac{T_{\beta}}{26} \qquad (3.1)$$

$$P_{\gamma} = \frac{T_{\gamma}}{26} \qquad (3.2)$$

where, P_{β} is the probability of an empty space being occupied by a fibrosis, T_{β} is the total number of fibrosis surrounding the empty space, P_{γ} is the probability of an empty space being occupied by an inflammatory cell, T_{γ} is the total number of inflammatory cells surrounding the empty space and 26 is the three-dimensional Moore neighborhood. The other types of autonomous agents, i.e., cardiomyocytes, TNF-α, *T. cruzi*, and bone marrow stem cells have a random concentration distribution. This concentration distribution is based on the experimental data obtained by Soares et al. (2004).

3.2.6. Input

The two external inputs are the initial fraction of fibrosis (0.22) and the initial fraction of inflammatory cells (0.08). These experimental values were previously obtained by Soares et al. (2004).

3.2.7. Submodels

The transition rules of states depend on a set of rules that establish the state of all sites. This computational model simulates the stem cell properties of apoptosis and differentiation. Our model

has five transition rules for simulating the chagasic tissue regeneration after bone marrow stem cell transplantation. The transition rules are similar to our previous two-dimensional agent-based model. However, we have changed the rule for increase in fibrosis. A schematic representation to explain the transition rules of our model is shown in Fig. 3.1. The rules are as follows:

• In the neighborhood of an inflammatory cell, if the number of bone marrow stem cells is different from zero, the inflammatory cell site changes into an empty space (Soares et al., 2004). This rule simulates the occurrence of apoptosis in an inflammatory cell.

• In the neighborhood of a fibrosis, if the number of bone marrow stem cells is different from zero, the fibrosis site changes into an empty space (Soares et al., 2004). This rule simulates the fibrosis regression in the heart.

• In the neighborhood of an empty space site, if there is a cardiomyocyte and the number of bone marrow stem cells is greater than a determinate value, the empty space site changes into cardiomyocyte (Orlic et al., 2001). This rule simulates the differentiation of bone marrow stem cells into cardiomyocyte.

• In the neighborhood of an empty space site, if the number of inflammatory cells and *T. cruzi* is different from zero, the empty space site changes into an inflammatory cell (Higuchi et al., 2003). This rule simulates the inflammatory cell migration to infected tissue.

• In the neighborhood of an empty space site, if the number of TNF-α, fibrosis (Soares et al., 2004) and inflammatory cell (Higuchi et al., 2003) is different from zero, the empty space site changes into a fibrosis. This rule simulates the fibrosis growth in the neighborhood of a fibrosis already established.

3.3. Results and discussion

For each simulation 20 runs were performed to estimate the parameters that describe the chronic chagasic cardiomyopathy regeneration after bone marrow stem cell transplantation. The optimization of the parameters was done using the fit with the experimental data. In our model, simulations run for 1000 time steps correspond to the experimental data (Soares et al., 2004) of five months of the real system. The data of inflammatory cells and fibrosis were normalized to permit a comparison between computational and experimental results. The normalization of the

computational and experimental data was done by the same method. All data were divided by the initial value of the respective data set.

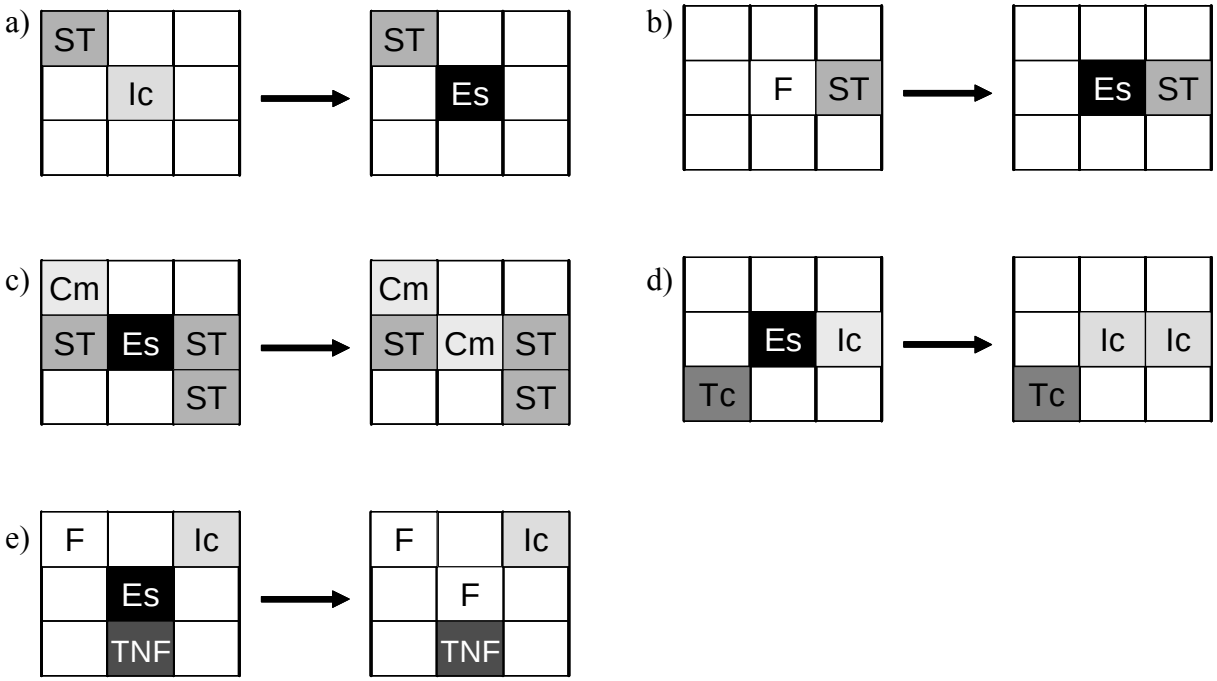


Fig. 3.1. Schematic representation of the transition rules in our model. The Moore neighborhood in the transition rules are represented in two-dimensions for clarity. However, in the model we use the three-dimensional Moore neighborhood. (a) In the neighborhood of an inflammatory cell (Ic), if the number of bone marrow stem cells (ST) is different from zero, the inflammatory cell site changes into empty space (Es) (Soares et al., 2004). (b) In the neighborhood of a fibrosis (F), if the number of ST is different from zero, the F site changes into Es (Soares et al., 2004). (c) In the neighborhood of an Es, if there is a cardiomyocyte (Cm) and the number of ST is greater than a certain value, the empty space site changes into Cm (Orlic et al., 2001). (d) In the neighborhood of an Es, if the number of Ic and *T. cruzi* (Tc) is different from zero, the Es site changes into an Ic (Higuchi et al., 2003). (e) In the neighborhood of an Es, if the number of TNF-α (TNF), F (Soares et al., 2004) and Ic (Higuchi et al., 2003) is different from zero, the empty space site changes into a F.

In Fig. 3.2 we compare both the three-dimensional computational model and the two-dimensional one by using previous (Galvão et al., 2008) and new rules. With the previous rules, the influence of inflammatory cells in the fibrosis progression was not included. The experimental results were obtained by Soares et al. (2004). In this experiment, 8-week-old female or male BALB/c mice were infected with *T. cruzi* trypomastigotes and treated 6 month later with adult bone marrow cells by the intravenous route. These types of cells contain bone marrow stem cells and many other types of cells. Groups of chronic chagasic mice treated with bone marrow cells were sacrificed 1, 2, 3, 5 and 6 months after treatment. At this time points, the percentage of fibrosis was determined and the number of inflammatory cells per square millimeter was counted in the heart.

We can observe that the variability of the extended model is smaller than our previous model. Also, the comparison of the extended and the earlier model shows that the curves of fibrosis and inflammatory cells have a different behavior. In the three-dimensional model, the quantity of

fibrosis is more similar to the experimental results when a new rule for fibrogenesis is taken into account (Soares et al., 2004). In the three-dimensional model by using the previous rules, the quantity of fibrosis does not decrease and rapidly stabilize due to the large fraction of TNF-α in the beginning of the simulation. In the two-dimensional model by using the new rules, the quantity of fibrosis decreases very rapidly and stabilizes near zero due to the small fraction of TNF-α. In this way, we can verify that the kinetics of fibrosis is related to the initial fraction of TNF-α and a better comparison of experimental and computational results is due to the new rule for the development of fibrosis.

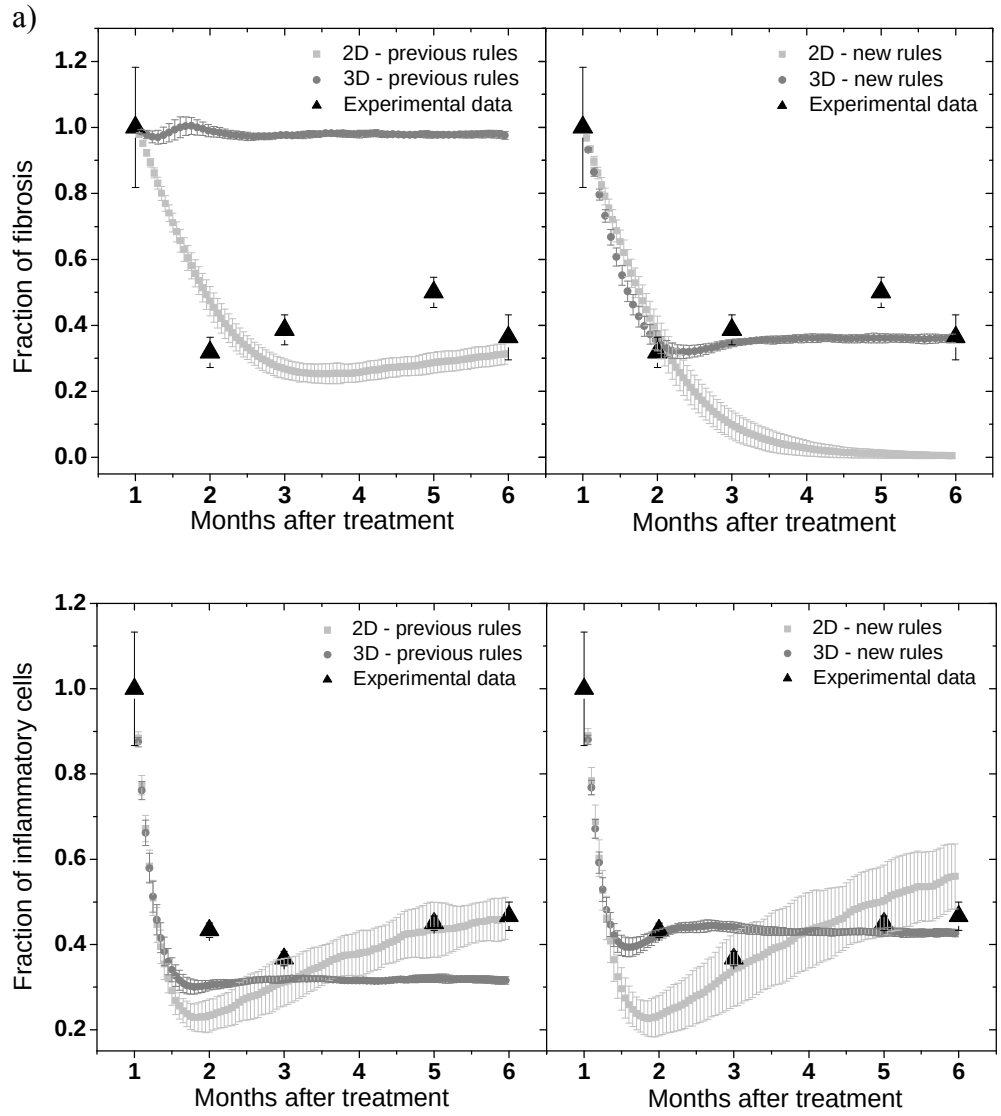


Fig. 3.2. Comparison of the three-dimensional computational model with the two-dimensional one by using previous (Ref. [28]) and new rules. In the three-dimensional model, the parameters used were the initial fraction of the lattice-sites occupied (So) equaling 0.8, the initial fraction of TNF-α equaling 1.1 x 10-3, the initial fraction of *T. cruzi* (Tc) equaling 1.2 x 10-4, the initial fraction of bone marrow stem cells (Bmst) equaling 2.5 x 10-3 and the number of bone marrow stem cells required for the differentiation of a bone marrow stem cell into a cardiomyocyte (Dcm) to occur is greater than 2. In the two-dimensional computational model [28], the parameters used were So = 0.8, Tc = 5.5 x 10-4, TNF-α = 4.5 x 10-3, Bmst = 10-2, and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Fibrosis regression. (b) Apoptosis of inflammatory cells.

Additionally, the different behavior of the curve that quantifies the number of inflammatory cells is due to the three-dimensional Moore neighborhood. In this type of neighborhood, the movement of the agents is facilitated because each site has twenty six adjacent neighbors and can move in three directions. In the two-dimensional Moore neighborhood, each site has eight adjacent neighbors and can move only in two directions. Therefore, in this type of neighborhood the agents have a greater probability to stay confined within a local region. This implies a larger variability and a slower stabilization. Then, the number of inflammatory cells stabilizes more rapidly when the number of adjacent neighbors is larger. Thus, the simulation of the heart in three-dimensions and the inclusion of inflammatory cells in the fibrosis progression improve the reproduction of the chagasic cardiac tissue after bone marrow stem cell transplantation.

Fig. 3.3 shows the kinetics of fibrosis regression by using different concentration patterns for fibrosis and inflammatory cells. We can observe that if inflammatory cells and fibrosis are randomly placed the quantity of fibrosis first decrease less and then it does not increase. If inflammatory cells are randomly placed and fibrosis are placed in empty spaces in the neighborhood of a fibrosis with a certain probability, the quantity of fibrosis decreases a little more but, also, it does not increase at the end. However, if inflammatory cells and fibrosis are placed with the probability described in our computational model, the quantity of fibrosis first decreases, then increases and finally stabilizes. In this curve, the fibrosis decrease is more rapid because, as the inflammatory cells are more concentrated in the beginning of the simulation, the probability of an inflammatory cell meeting a fibrosis is smaller. The fibrosis stabilization occurs when the inflammatory cells have a homogeneous distribution in the tissue. This latter computational result has a fairly fit with the experimental data (Soares et al., 2004).

In this extended computational model, the clusterization of fibrosis and inflammatory cells through the simulation grid is consistent with the distribution in real systems. In our previous article (Galvão et al., 2008), the two-dimensional distribution was compared qualitatively to a picture of cardiac chagasic tissue (Soares et al., 2004). In those heart sections, the regions of fibrosis and inflammatory cells are larger and more concentrated in the non-transplanted mice. That picture is two-dimensional and, therefore, if many pictures of heart sections are taken and the image of three-dimensional heart is rebuilt, the areas of fibrosis and inflammatory cells should continue greater and more concentrated in the non-transplanted mice. Then, the pattern distribution of fibrosis and inflammatory cells in our three-dimensional model represents the chagasic cardiac tissue.

In this way, originally the areas of fibrosis and inflammatory cells are more concentrated, but they do not possess a uniform density. The evolution of this extended model simulates the dissolution of fibrosis areas and inflammatory cells. Also, it simulates the reduction in the quantity of fibrosis and in the number of inflammatory cells. The dissolution and reduction of fibrosis and

inflammatory cells also occur in the experimental model developed by Soares et al. (2004). Therefore, this model indicates that the clusterization of fibrosis and inflammatory cells in the beginning of the simulation has a major significance because it is responsible for the chagasic tissue regeneration after bone marrow stem cell transplantation.

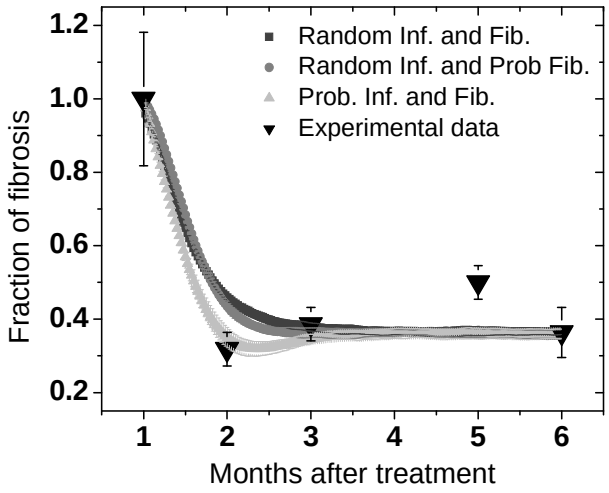


Fig. 3.3. Kinetics of fibrosis regression by using different concentration distribution for fibrosis and inflammatory cells. In the darkest grey curve, inflammatory cells and fibrosis are randomly placed. In the grey curve, all inflammatory cells and some fibrosis are randomly placed. The other fibrosis are placed in empty spaces in the neighborhood of a fibrosis with a certain probability. In the light grey, some fibrosis are randomly placed and the others are placed in empty spaces in the neighborhood of a fibrosis with a certain probability. The same occurs to the inflammatory cells. The parameters used were So = 0:8, TNF- = 1.1 x 10^{-3}, Tc = 1.2 x 10^{-4}, Bmst = 2.5 x 10^{-3}, and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (See Fig. 3.2 for the meaning of the labels.)

The experimental curve of fibrosis (Soares et al., 2004) increases in the third month after treatment. This probably occurs because there are inflammatory cells and TNF-α in the heart mice. Consequently, the expansion of fibrosis persists. In the fourth month after treatment, the experimental curve of inflammatory cells (Soares et al., 2004) shows an increase that is possibly caused by the presence of *T. cruzi* in the cardiac tissue of mice. In Fig. 3.4 we show the kinetics of fibrosis regression and apoptosis of inflammatory cells for different fractions of lattice-site occupation. Our results show that the larger the number of sites occupied, the smaller the final fraction of fibrosis or inflammatory cells. This occurs because it is more difficult for an inflammatory cell and a TNF-α to meet a fibrosis once the fraction of empty spaces is small. In addition, it is more difficult for a *T. cruzi* to meet an inflammatory cell when the fraction of empty spaces is small.

Fig. 3.5 shows the variation of the initial fraction of bone marrow stem cells for fibrosis regression and apoptosis of inflammatory cells. The results for fibrosis increase because our computational model possesses inflammatory cells and TNF-α. Similarly, our results for inflammatory cells increase because it possesses *T. cruzi*. We can observe that the behavior of

fibrosis regression is more affected by a small difference in the initial fraction of the bone marrow stem cells than the apoptosis behavior of inflammatory cells. In this model, the quantity of fibrosis and the number of inflammatory cells stabilize during the time evolution.

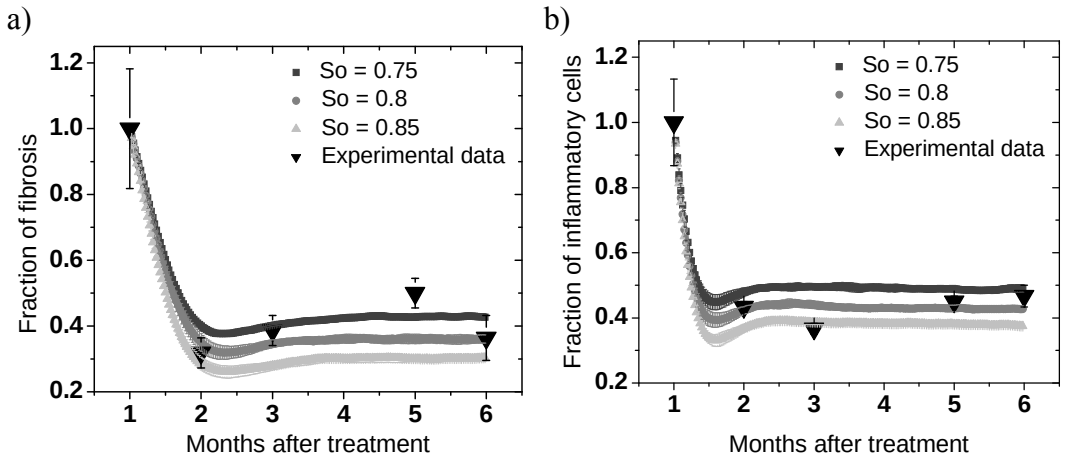


Fig. 3.4. Chronic chagasic cardiomyopathy regeneration by using different fractions of lattice-site occupation. The other parameters used were TNF-α = 1.1 x 10^{-3}, Tc = 1.2 x 10^{-4}, Bmst = 2.5 x 10^{-3}, and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Fibrosis regression. (b) Apoptosis of inflammatory cells. (See Fig. 3.2 for the meaning of the labels.)

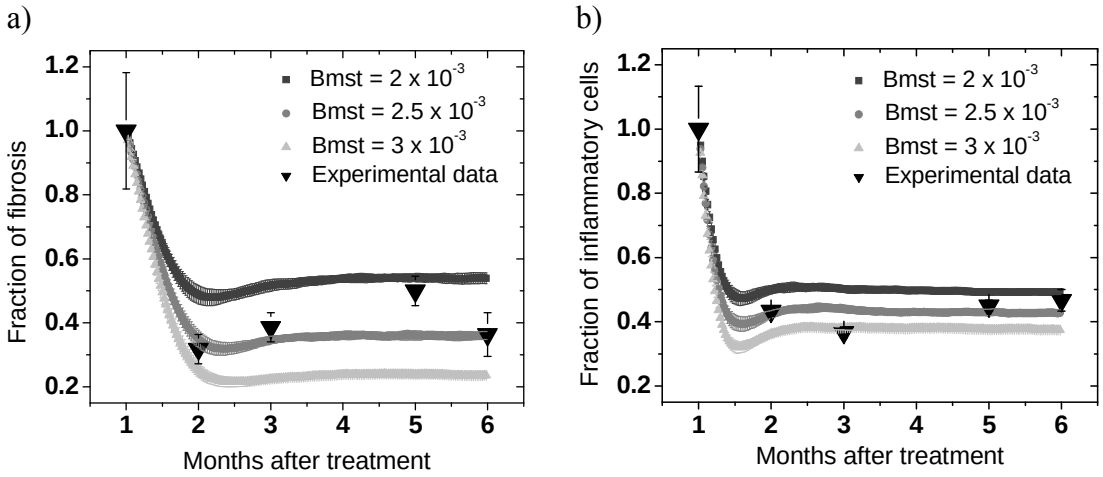


Fig. 3.5. Chronic chagasic cardiomyopathy regeneration by using different fractions of bone marrow stem cells. The other parameters used were So = 0.8, TNF-α = 1.1 x 10^{-3}, Tc = 1.2 x 10^{-4}, and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Fibrosis regression. (b) Apoptosis of inflammatory cells. (See Fig. 3.2 for the meaning of the labels).

In Fig. 3.6(a) we vary the initial fraction of TNF-α and in Fig. 3.6(b) the initial fraction of *T. cruzi*. Our data show that a small difference in the initial fraction of TNF- modifies the increase in the quantity of fibrosis. In the same way, a diminutive difference in the initial fraction of *T. cruzi* modifies the increase in the number of inflammatory cells. Hence, the increase in the quantity of fibrosis depends on the initial value of TNF-α and the increase in the number of inflammatory cells depends on the initial value of *T. cruzi*.

Fig. 3.7 shows the differentiation of bone marrow stem cells into cardiomyocyte. In our model, the increase in the number of new cardiomyocytes is linear. In the neighborhood of an empty space site, if there is a cardiomyocyte and the number of bone marrow stem cells is greater than 2, the increase in the total number of cardiomyocytes grows around 1.31%, and if the number of bone marrow stem cells is greater than 3, the number of cardiomyocytes grows around 0.02%. As the experimental results [4] do not quantify the number of new cardiomyocytes, we can only suggest that the increase of cardiomyocytes is less realistic when the number of bone marrow stem cells is greater than 3 because it is very small.

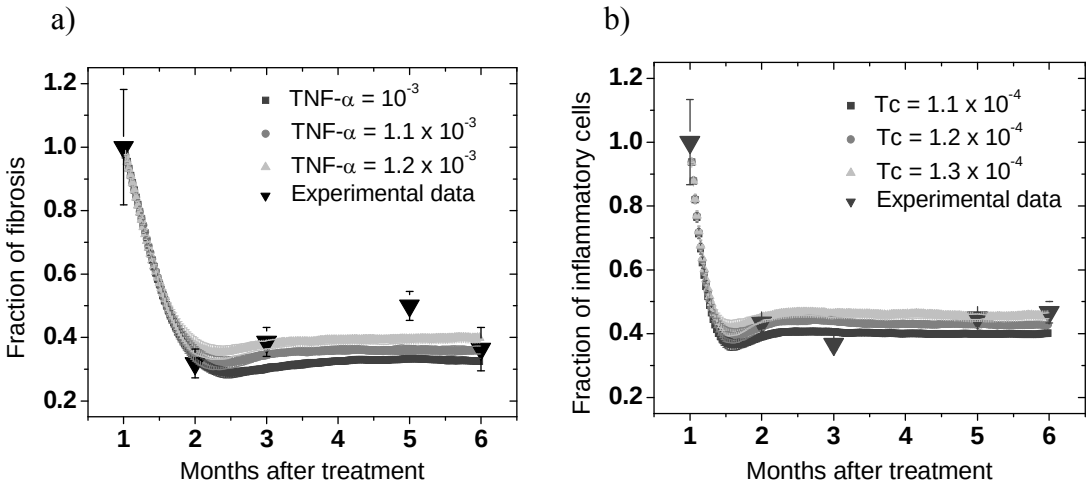


Fig. 3.6. Chronic chagasic cardiomyopathy regeneration by using different fractions of TNF-α and *T. cruzi*. The other parameters used were So = 0.8, Bmst = 2.5 x 10-3, and Dcm > 2. The experimental data were obtained by Soares et al. (2004). (a) Fibrosis regression. In this simulation we used Tc = 1.2 x 10-4. (b) Apoptosis of inflammatory cells. In this simulation we used TNF-α = 1.1 x 10-3. (See Fig. 3.2 for the meaning of the labels).

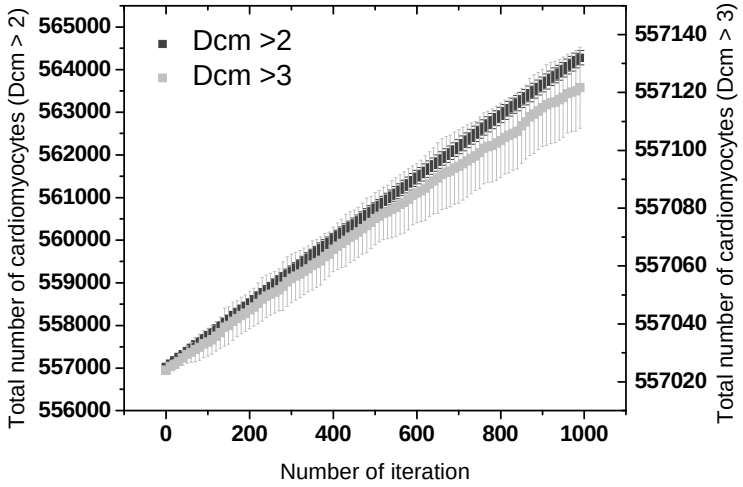


Fig. 3.7. Chronic chagasic cardiomyopathy regeneration by using different number of bone marrow stem cells required for differentiation of a bone marrow stem cell into a cardiomyocyte. The other parameters used were So = 0.8, Bmst = 2.5 x 10-3, TNF-α = 1.1 x 10-3, and Tc = 1.2 x 10-4. (See Fig. 3.2 for the meaning of the labels).

Thus, according to the experimental data the best computational parameters of our model to describe the chagasic tissue regeneration are the fraction of the number of sites occupied equaling 0.8, fraction of *T. cruzi* equaling 1.2 x 10^{-4}, fraction of TNF-α equaling 1.1 x 10^{-3}, fraction of bone marrow stem cells equaling 2.5 x 10^{-3}, and the number of bone marrow stem cells required to occur differentiation of a bone marrow stem cell into a cardiomyocyte greater than 2. The value of the computational parameters that provide the best fit to the experimental data was not compared to experimental measurements, because the value of these parameters was not determined in the experiment developed by Soares et al. (2004). This means that we can only infer that our best parameters are reasonable, for the reason that the quantity of bone marrow stem cells transplanted in chagasic mice was small (Soares et al., 2004) and various researches have emphasized that the quantity of *T. cruzi* in the chronic phase of Chagas' disease is disproportionately low (Higuchi et al., 2003).

3.4. Conclusion

In this paper, we have presented a three-dimensional multi-agent-based model for regeneration of the Chagas' disease after bone marrow stem cell transplantation. Our new results suggest that the inflammatory cells are important for the development of fibrosis. Therefore, the immune response can be relevant to the pathology of Chagas' disease. According to stem cell property, our model can simulate the apoptosis and differentiation. Bone marrow stem cells can promote apoptosis in inflammatory cells and can differentiate into cardiomyocytes. Also, bone marrow stem cells can promote fibrosis regression. In our computational model, the most important aim is to comprehend the contribution of different types of cells in the chronic chagasic cardiomyopathy regeneration. The principal hypothesis is based on the clusterization of fibrosis and inflammatory cells. This type of distribution allows an appropriate modeling of chagasic cardiac tissue regeneration.

Our computational data were compared with experimental data, excepting for the differentiation of bone marrow stem cells into cardiomyocytes, giving a fairly fit. The results suggest that an important factor to characterize the kinetics of chronic chagasic cardiomyopathy regeneration after bone marrow stem cell transplantation is the cellular clusterization. Also, the three-dimensional neighborhood can reproduce better the cellular kinetic in the chagasic tissue because the cells have less probability to stay confined within a local region. Additionally, the initial fraction of bone marrow stem cells and TNF-α modifies the regression rate of fibrosis. Lastly, the initial fraction of bone marrow stem cells and *T. cruzi* modifies the apoptosis rate of inflammatory cells.

Acknowledgements

This work was supported by the Brazilian agency FAPESB.

References

Adimy, M., Crauste, F., Abdllaoui, A. (2008a). Discrete-maturity structured model of cell differentiation with applications to acute myelogenous leukemia. Journal of Biol. Syst., 6, 395-424.

Adimy, M. et al. (2008b). Numerical integration of a mathematical model of hematopoietic stem cell dynamics. Comput. Math. Appl., 56, 594-606.

Agur, Z., Daniel, Y., Ginosar, Y. (2002). The universal properties of stem cells as pinpointed by a simple discrete model. J. Math. Biol. 44, 79-86.

Andrade, L.O., Andrews, N.W. (2005) The *Trypanosoma cruzi*–host-cell interplay: location, invasion, retention. Nature Rev. Microb., 3, 819-823.

Bianco, P. et al. (2001) Bone marrow stromal stem cells: nature, biology, and potential applications. Stem Cells, 19, 180–192.

de Almeida, M.C., Moreira, H.N. (2007) A mathematical model of immune response in cutaneous Leishmaniasis. J. Biol. Syst., 15, 313-354.

d'Inverno, M., Saunders, R. (2005) Agent-based modelling of stem cell organisation in a niche, Engineering self-organising systems: methodologies and applications – LNCS, 3464, 52-68.

Ferrer, J. et al. (2007) Individual-based model and simulation of *Plasmodium falciparum* infected erythrocyte in vitro cultures. J. Theor. Biol., 248, 448-459.

Galvão, V., Miranda, J.G.V., Ribeiro-dos-Santos, R. (2008) Development of a two-dimensional agent-based model for chronic chagasic cardiomyopathy after stem cell transplantation. Bioinformatics, 24, 2051-2056.

Grimm, V. et al. (2006). A standard protocol for describing individual-based and agent-based models. Ecol. Model., 198, 115 - 126.

Grove, J.E., Bruscia, E., Krause, D.S. (2004) Plasticity of bone marrow-derived stem cells. Stem Cells, 22, 487-500.

Higuchi, M.L. et al. (2003) Pathophysiology of the heart in Chagas' disease: current status and new developments. Cardiovasc. Res., 60, 96-107.

Isasi, S.C., Sibona, G.J., Condat, C.A. (2001) A simple model for the interaction between *T. cruzi* and its antibodies during Chagas infection. J. Theor. Biol., 208, 1-13.

Joung, J.-G. et al. (2006) Identification of regulatory modules by co-clustering latent variable models: stem cell differentiation. Bioinformatics, 22, 2005-2011.

Nagarajan, R., Aubin, J.E., Perterson, C.A. (2005) Robust dependencies and structures in stem cell differentiation. Int. J. Bifurcat. Chaos, 15, 1503-1514.

Nelson, P., Velasco-Hernández, J.X. (2002) Modeling the immune response to parasitic infections: Leishmaniasis and Chagas disease. Comments Theor. Biol., 6, 161.

Orlic, D. et al. (2001) Bone marrow cells regenerate infarcted myocardium. Nature, 410, 701–705.

Pisu, M., Concas, A. and Cao, G. (2007) A novel simulation model for stem cells differentiation. J. Biotechnol., 130, 171-182.

Pisu, M. et al (2008) A simulation model for stem cells differentiation into specialized cells of non-connective tissues. Comput. Biol. Chem., 32, 338-344.

Sahimi, M., Mehrabi, A.R., Naeim, F. (1997) Discrete stochastic model for self-renewal and differentiation of progenitor cells. Phys. Rev. E, 55, R2111.

Sell, S. (2002) Stem Cell Handbook, Humana Press, Totowa, NJ, 2002.

Sibona, G.J., Condat, C.A. (2002) Dynamic analysis of a parasite population model. Phys. Rev. E, 65, 031918.

Sibona, G.J., Condat, C.A., Isasi, C. (2005) Dynamics of the antibody-*T. cruzi* competition during Chagas infection: prognostic relevance of intracellular replication. Phys. Rev. E, 71, 020901.

Soares, M.B.P., Pontes-de-Carvalho, L., Ribeiro-dos-Santos, R. (2001) The pathogenesis of Chagas' disease: when autoimmune and parasite-specific immune responses meet. An. Acad. Bras. Cienc., 73, 547-559.

Soares, M.P.B. et al. (2004) Transplanted bone marrow cells repair heart tissue and reduce myocarditis in chronic chagasic mice. Am. J. Pathol., 164, 441-447.

Takamatsu, A., Fujii, T., Endo, I. (2000) Time Delay Effect in a Living Coupled Oscillator System with the *Plasmodium* of *Physarum polycephalum*. Phys. Rev. Lett., 85, 2026-2029.

Tannenbaum, E. et al. (2005) Evolutionary dynamics of adult stem cells: comparison of random and immortal-strand segregation mechanisms. Phys. Rev. E, 71, 041914.

Wolfram, S. (1986) Theory and Applications of Cellular Automata, World Scientific, Singapore.

Woolf, P.J. et al. (2005) Bayesian analysis of signaling networks governing embryonic stem cell fate decisions. Bioinformatics, 21, 741-753.

Zhdanov, V.P. (2008) Simulation of proliferation and differentiation of cells in a stem-cell niche. Physica A, 387, 6126-6136.

Zorzenon dos Santos, R.M. et al. (2007) On the study of the dynamical aspects of parasitemia in the blood cycle of malaria. Eur. Phys. J. Special Topics, 143, 125-132.

CHAPTER 4

A three-dimensional multi-agent-based model for evolution of Chagas' disease

Abstract

A better understanding of Chagas' disease is important because the knowledge about the progression and the participation of the different types of cells in this disease are still lacking. To clarify this system, the kinetics of inflammatory cells and parasite nests was shown in an experiment. Using this experimental data, we have developed a three-dimensional multi-agent-based computational model for the evolution of Chagas' disease. Our model includes five different types of agents: inflammatory cell, fibrosis, cardiomyocyte, fibroblast, and *T. cruzi*. Fibrosis is fixed and the other types of agents can move through the empty space. They move randomly by using the Moore neighborhood. This model reproduces the acute and chronic phases of Chagas' disease and the volume occupied by all different types of cells in the cardiac tissue.

Keywords: computational model, chronic chagasic cardiomyopathy, autonomous agent, parasitemia.

4.1. Introduction

Chagas' disease, caused by the protozoan *Trypanosoma cruzi*, affects tens of millions of people worldwide (WHO, 2002). *T. cruzi* parasite is transmitted by a blood-sucking insect of the sub-family Triatominae, or by blood transfusion. Chagas' disease has two distinct phases of progression. The acute phase is characterized by high numbers of *T. cruzi*, intense parasitism and inflammation. In the chronic phase, most of the *T. cruzi*-infected people remain in the asymptomatic or indeterminate form without any signs or clinical symptoms. However, around 30% of them present cardiac complications. Chronic chagasic cardiomyopathy (CChC) is characterized by the

presence of inflammatory infiltrates in the heart, interstitial fibrosis, and scarce quantity of *T. cruzi* parasites (Higuchi et al., 2003; Soares et al., 2001a, 2004; Teixeira et al., 2006; Coura, 2007). In this phase, the occurrence of *T. cruzi* is related with a severe or moderate inflammation (Higuchi et al., 2003). The development of fibrosis is caused by the proliferation of fibroblasts and the subsequent deposition of interstitial collagens (Ten Tusscher and Panfilov, 2007).

The complex life cycle of *T. cruzi* has three morphological stages (trypomastigote, amastigote and epimastigote). An infected insect vector eliminates metacyclic trypomastigotes with the feces during the blood meal. Metacyclic trypomastigote enters into the bloodstream through the bite wound or through intact mucosal membranes. This form of *T. cruzi* can invade different types of cells, including macrophages, heart muscle, nerve tissue and digestive tract cells. Inside the cell, trypomastigotes differentiate into intracellular amastigotes. Amastigotes multiply by binary fission and after some divisions, amastigotes transform into trypomastigotes that are released after lyses of the host cell. The circulate parasites can invade other cells or be taken up by the insect vector during a blood meal. In the insect midgut, trypomastigotes transform into epimastigotes, which multiply and migrate to the hindgut. In the hindgut, they differentiate back to metacyclic trypomastigotes and are released in the insect feces (Andrade and Andrews, 2005; Kelly, 2000).

Recently, different theoretical models have been proposed to investigate the kinetics of diseases caused by parasites. Examples include aids (Bailey et al., 1992; Corne and Frisco, 2008; Figueirêdo et al., 2008), malaria (Ferrer et al., 2007; Hoschen et al., 2000; McKenzie and Bossert, 2005; Zorzenon dos Santos et al., 2007), leishmaniasis (Nelson and Velasco-Hernández, 2002; Almeida and Moreira, 2007), toxoplasmosis (González-Parra et al., 2009; Kafsack et al., 2007), tuberculosis (Segovia-Juarez et al., 2004; Magombedze et al, 2006) and Chagas' disease (Galvão et al., 2008; Galvão and Miranda, Isasi et al., 2001; Nelson and Velasco-Hernández, 2002; Sibona and Condat, 2002; Sibona et al., 2005). The inclusion of inflammatory cell, fibrosis, cardiomyocyte and fibroblast is lacking in the evolution models for the Chagas' disease. Thus, we propose in this article a three-dimensional multi-agent-based model for the evolution of Chagas' disease including these types of cells.

4.2. Computational model

The description of our computational model follows the standard ODD protocol (Overview, Design concepts, and Details) for individual-based and agent-based models (Grimm *et al.*, 2006).

4.2.1. Purpose

The purpose of this computational model is to understand the participation of different types of cells in the development of Chagas' disease.

4.2.2. State variables and scales

We have developed a software program in C++ language to simulate the evolution of Chagas' disease. A three-dimensional lattice consisting of a grid with 100 x 100 x 100 points in size and periodic boundary conditions is employed to represent the cardiac tissue. The biological unit of the lattice simulated is obtained by the comparison with the experiment (Soares et al., 2001b). Our computational model includes different types of agents in which each type corresponds to a different type of cell. Each lattice site represents the space region that can only be occupied by a type of cell. The agents correspond to cellular groups with a similar pattern of behavior because the number of cells in the myocardial tissue is large. In a site, if the number of agent is zero; the site is referred to as empty space. All actions occur at constant intervals called time steps and the description of the chagasic tissue is discrete both in time and space.

This computational model includes five different types of agents: inflammatory cell, fibrosis, cardiomyocyte, fibroblast, and *T. cruzi*. The parameters are the total number of agents, the initial fraction of inflammatory cells, and the initial fraction of fibroblast. The fraction of cardiomyocytes is given by the difference between the total number of agents and the total number of inflammatory cells, fibroblast and *T. cruzi*. The experimental data (Soares et al., 2001b) of 7 months correspond to the 80 time steps of the model; therefore, each time step corresponds to around 2.625 days of experiment.

4.2.3. Process overview and scheduling

Inflammatory cells, cardiomyocytes, fibroblast, and *T. cruzi* can move through the empty spaces. For simulating the cellular movement, each type of agent is selected randomly and can jump to empty space. These types of agent move randomly by using the three-dimensional Moore neighborhood. In this type of neighborhood, each cell possesses 26 neighboring cells situated in positions above, below, and to the sides (Wolfram, 1986). Fibrosis is fixed, i.e., this type of agent can not move.

4.2.4. Design concepts

Emergence: Population kinetics and structure emerge from the behavior of different types of cells. The evolution of chagasic tissue is completely represented by local and deterministic rules.
Sensing: Agents know their state, they recognize the state of the other agents and they apply the transition rules according to the state of their first neighbors.
Interaction: The system evolution is determined by interactions between the different types of agents (inflammatory cell, fibrosis, cardiomyocyte, fibroblast, and *T. cruzi*). The interactions occur in the neighborhood of an agent, i.e., they are locals.
Observation: The number of inflammatory cells, parasite nests and quantity of fibrosis are statistically investigated at each time step.

4.2.5. Initialization

Inflammatory cells, cardiomyocytes, fibroblast, and *T. cruzi* have a random distribution. This distribution pattern is based on the experimental data obtained by Soares et al. (2001b). Initially, the cardiac tissue does not possess fibrosis; this type of agent only appears in the time evolution.

4.2.6. Input

We employ the same infective inoculum of 100 *T. cruzi* used by Soares et al. (2001b). Therefore, we use the same quantity of *T. cruzi* in all simulations and we only have one external input.

4.2.7. Submodels

The time evolution is run in the complete lattice and each site changes its state according to a local deterministic rule, which depends only on the neighboring cells. The state of the all sites is determined through the transition rules that depend on a set of rules. This model simulates the acute and chronic phases of Chagas' disease using the cell properties in the chagasic tissue. Our model has eight transition rules for simulate the evolution of Chagas' disease. A schematic representation to explain the transition rules of our model is shown in Fig. 4.1. The transition rules are as follow:

(i) In the neighborhood of an empty space site, if the number of inflammatory cells and *T. cruzi* is different from zero, the empty space site changes into inflammatory cell (Higuchi et al., 2003). This rule simulates the migration of an inflammatory cell to infected tissue.

(ii) In the neighborhood of an inflammatory cell, if the number of *T. cruzi* is different from zero, the inflammatory cell site changes into empty space (Soares et al., 2001b). This rule simulates the inflammatory cell death by phagocytosis of *T. cruzi*.

(iii) In the neighborhood of a *T. cruzi*, if the number of inflammatory cells is different from zero, the *T. cruzi* site changes into empty space (Soares et al., 2001b). This rule simulates the *T. cruzi* phagocytosis by inflammatory cells.

(iv) In the neighborhood of an inflammatory cell, if the number of inflammatory cells is greater than a determined value, death of inflammatory cells occurs (Andrade, 1999), i.e., the inflammatory cell site changes into empty space. This rule simulates the inflammatory cells death by space and resource competition.

(v) In the neighborhood of an empty space, if the number of *T. cruzi* and cardiomyocytes is different from zero, the empty space site changes into *T. cruzi* (Andrade, 1999). *T. cruzi* replication occurs inside the cardiomyocyte, but our model cannot exactly simulate this replication, because the site can only be occupied by a different type of agent. Therefore, *T. cruzi* replication is simulated in the neighborhood of a cardiomyocyte.

(vi) In the neighborhood of *T. cruzi*, if the number of *T. cruzi* and cardiomyocytes is different from zero (de Souza et al., 2003), death of *T. cruzi* occurs, i.e., the *T. cruzi* site changes into empty space. The death of *T. cruzi* occurs inside the cardiomyocyte, but our model cannot exactly simulate this death, because the site can only be occupied by a different type of agent. Hence, the death of *T. cruzi* by resource and space competition is simulated in the neighborhood of a cardiomyocyte and it occurs more frequently in the acute phase of Chagas' disease.

(vii) In the neighborhood of a cardiomyocyte, if the number of inflammatory cells, *T. cruzi* (Andrade, 1999) and fibroblast (Ten Tusscher and Panfilov, 2007) is different from zero, the cardiomyocyte site changes into a new fibrosis. This rule simulates the formation of new fibrosis areas in the chagasic heart.

(viii) In the neighborhood of a cardiomyocyte, if the number of inflammatory cells (Andrade, 1999), fibrosis and fibroblast (Ten Tusscher and Panfilov, 2007) is different from zero, the cardiomyocyte site changes into fibrosis. This rule simulates the cluster formation of fibrosis by the auto-immune response.

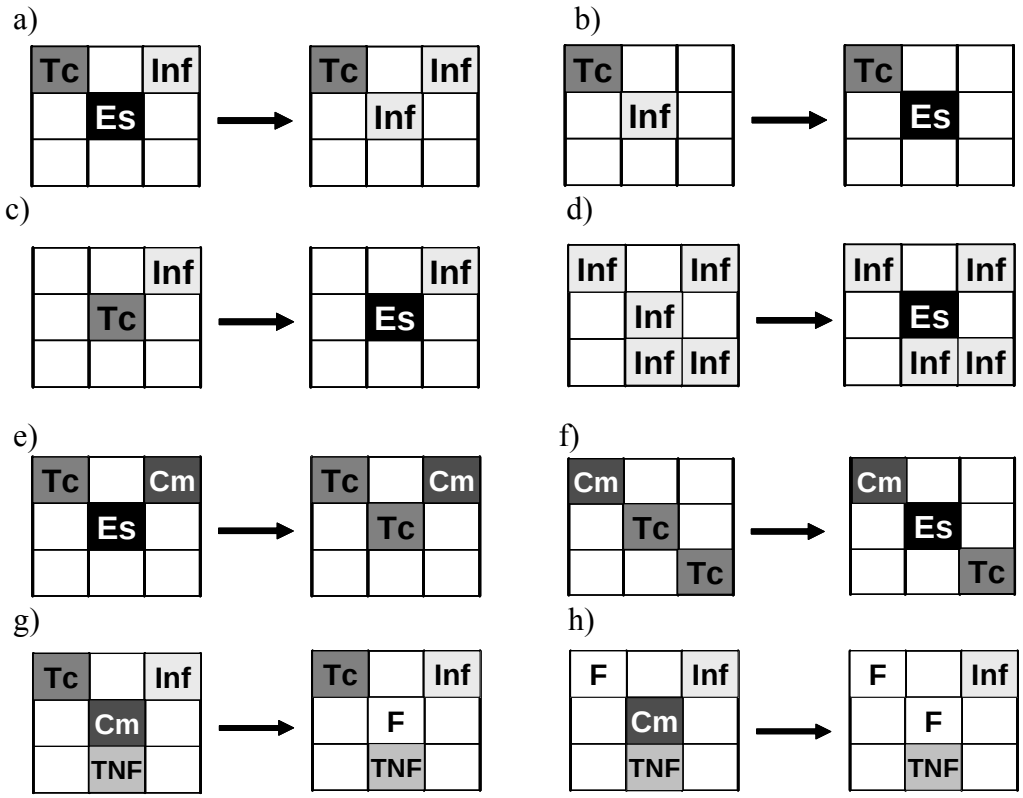


Fig. 4.1. Schematic representation of the transition rules. The Moore neighborhood is represented in two-dimensions for clarity. However, in our model we use the three-dimensional Moore neighborhood. (a) In the neighborhood of an empty space (Es), if the number of *T. cruzi* (Tc) and inflammatory cell (Ic) is different from zero, the Es site changes into Ic (Higuchi et al., 2003). (b) In the neighborhood of an Ic, if the number of Tc is different from zero, the Ic site changes into Es (Soares et al., 2001b). (c) In the neighborhood of a Tc, if the number of Ic is different from zero, the Tc site changes into Es (Soares et al., 2001b). (d) In the neighborhood of an Ic, if the number of Ic is greater than a determined value, the Ic site changes into Es (Andrade, 1999). (e) In the neighborhood of an Es, if the number of Tc and cardiomyocytes (Cm) is different from zero, the Es site changes into Tc (Andrade, 1999). (f) In the neighborhood of Tc, if the number of Tc and Cm is different from zero (de Souza et al., 2003), the Tc site changes into Es. (g) In the neighborhood of a Cm, if the number of Ic, Tc (Andrade, 1999) and TNF-α (TNF) (Soares et al., 2004) is different from zero, the Cm site changes into fibrosis (F). (h) In the neighborhood of a Cm, if the number of Ic (Andrade, 1999), F and TNF (Soares et al., 2004) is different from zero, the Cm site changes into F.

4.3. Results

For each set of parameters, the average value of 20 simulation runs was taken in order to better describe the development of Chagas' disease. The parameters optimization was carried out using the fit corresponds to the experimental data (Soares et al., 2001b) of seven months. The model steps were equally followed in all simulations. The inflammatory cells and parasite nests were

normalized to allow a comparison between computational and experimental data. The normalization was done by the same method. All data were divided by the larger value of the respective data set.

We compare the predictions of our multi-agent-based model with the experimental results for the evolution of Chagas' disease until the CChC obtained by Soares et al. (2001b). In this model, IL-4-deficient and wild-type BALB/c mice were infected by inoculation of *T. cruzi* by intraperitoneal route. Groups of mice were sacrificed at 0.66, 1, 1.33, 3, 4 and 7 months after infection. At these time points, the number of parasite nests per square centimeter and inflammatory cells per square millimeter were counted in the heart. After the acute phase, the authors noted that the hearts of wild-type mice had scarce inflammatory foci and the IL-4-deficient mice had scarce inflammatory foci and the IL-4 deficient mice had a visible inflammatory reaction. In addition, in the chronic phase the quantity of parasite nests is very small in both types of mice. We have compared our results with the experimental results of IL-4-deficient mice because these mice are more resistant to *T. cruzi* infection, the myocarditis is exacerbated, and they survive longer than the wild-type.

Figure 4.2 shows the kinetics of inflammatory cells and parasite nests for different fractions of lattice-site occupation. As our model is three-dimensional, this fraction corresponds to the volume fraction occupied by cells. The best fit of our data with the experimental data was obtained with the fraction of lattice-site occupation 0.7. According to Vinnakota and Bassingthwaighte (2004) the volume fraction occupied by all types of cells in the rat myocardial tissue is 0.694 ± 0.025. Hence, our model reproduces the cellular volume of the cardiac tissue. Additionally, our results show that the larger the number of sites occupied, the larger the final fraction of inflammatory cells. This happens because it is more difficult for an inflammatory cell to meet others inflammatory cells once the fraction of empty spaces is small. Also, the quantity of parasite nests spends a little more time to stabilize around zero because it is more difficult for a *T. cruzi* to meet an inflammatory cell when the fraction of empty spaces is small.

The kinetics of inflammatory cells and parasite nests for different fractions of inflammatory cells is shown in Fig. 4.3. We can observe that the number of inflammatory cells increases rapidly, after it has a gradual reduction and lastly tends to stabilize. The quantity of this type of cell continues decreasing when the number of *T. cruzi* is practically zero due to the occurrence of inflammatory cell death by space competition [rule (iv)]. The number of parasite nests increases very rapidly, after decreases, and lastly is around zero. In the acute phase of Chagas' disease, the number of inflammatory cells and parasite nests is large. In the chronic phase, the number of inflammatory cells tends to become stable and the number of parasite nests is scarce. In this way, our model simulates the acute and chronic phases of this disease.

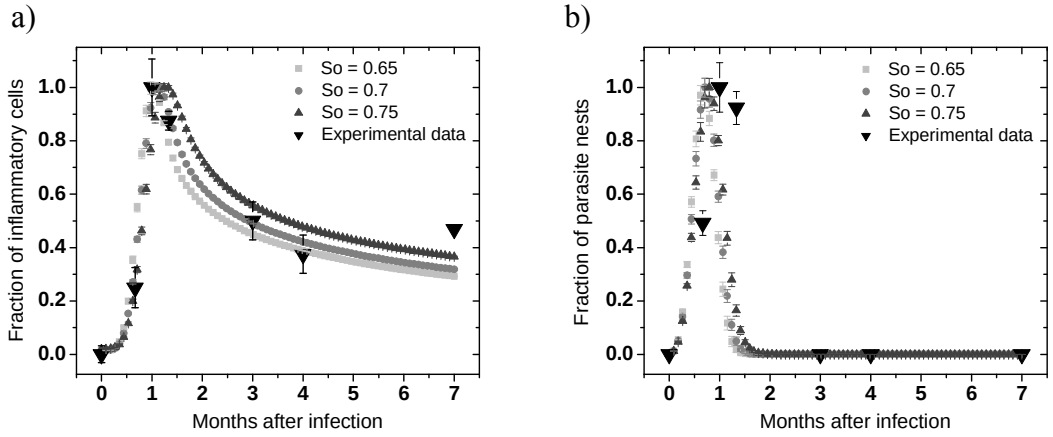


Figure 4.2. Evolution of Chagas' disease by using different fractions of lattice-site occupation (So). The other parameters used were the initial fraction of inflammatory cells (Inf) equaling 2 x 10^{-3}, initial fraction of TNF-α equaling 2 x 10^{-3} and the number of inflammatory cells required for apoptosis of an inflammatory cell (NInf) is greater than 3. The experimental data were obtained by Soares *et al.* (2001b). (a) Kinetics of inflammatory cells. (b) Kinetics of parasite nests.

In figure 4.4 we vary the number of inflammatory cells required for death of an inflammatory cell. Our results show that the difference of one unit in the number of inflammatory cells modifies the behavior of the inflammatory cells. The number of inflammatory cells decreases very rapid when the number of inflammatory cells required for death of an inflammatory cell is greater than 2 and this cause an increase in the number of parasites. Consequently, the number of inflammatory cells increases because the number of parasites is elevated. After this, the number of inflammatory cells and parasites tends to stabilize in a high value. However, the curve behavior of the parasite is less affected by the variation of this parameter because it does not depend directly on this parameter for the transition rule.

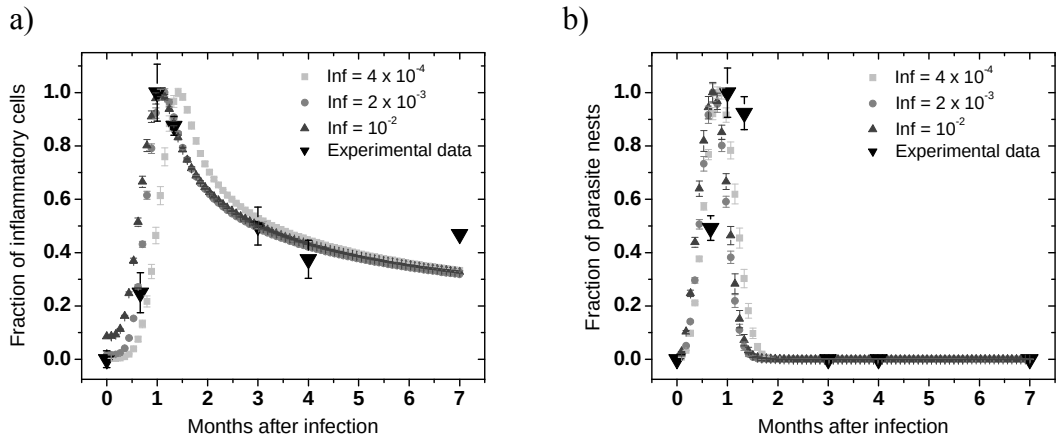


Fig. 4.3. Evolution of Chagas' disease by using different fractions of inflammatory cells. The other parameters used were So = 0.7, TNF-α = 2 x 10^{-3} and NInf > 3. The experimental data were obtained by Soares *et al.* (2001b). (a) Kinetics of inflammatory cells. (b) Kinetics of parasite nests. (See Fig. 4.2 for the meaning of the labels.)

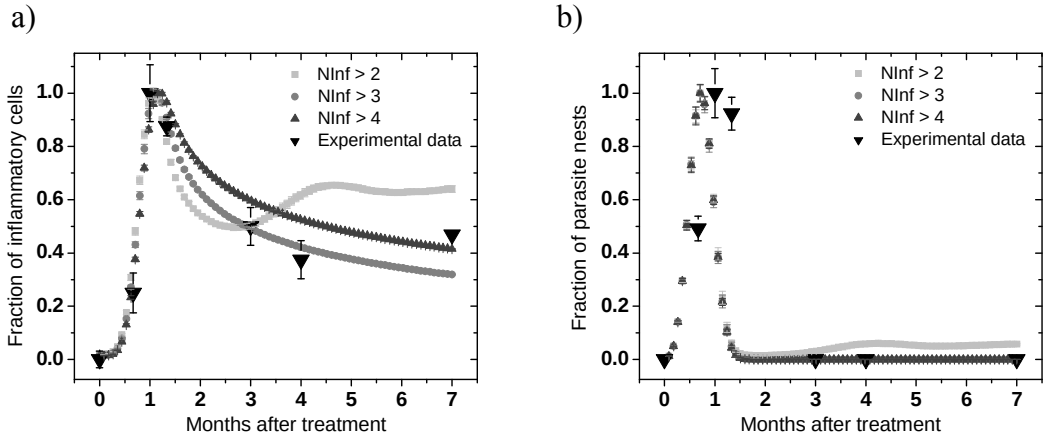


Fig. 4.4. Evolution of Chagas' disease by using different number of inflammatory cells required for apoptosis of an inflammatory cell. The other parameters used were So = 0.7, Inf = 2 x 10^{-3} and TNF-α = 2 x 10^{-3}. The experimental data were obtained by Soares *et al.* (2001b). (a) Kinetics of inflammatory cells. (b) Kinetics of parasite nests. (See Fig. 4.2 for the meaning of the labels.)

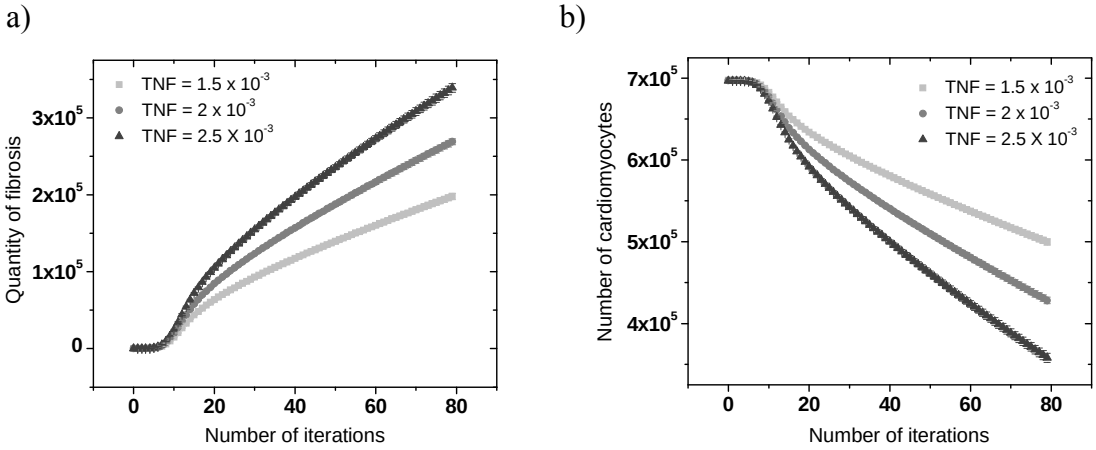


Fig. 4.5. Evolution of Chagas' disease by using different fractions of TNF-α. The other parameters used were So = 0.7, Inf = 2 x 10^{-3} and NInf > 3. (a) Kinetics of fibrosis progression. (b) Kinetics if cardiomyocytes apoptosis. (See Fig. 2 for the meaning of the labels.)

The kinetics of fibrosis progression and cardiomyocytes death for different fractions of fibroblasts is shown in Figure 5. Our data show that the larger the final fraction of fibrosis, the smaller the final fraction of cardiomyocytes. This is related because several studies on CChC suggest that during the expansion of this disease the fibrosis progression is associated with the cardiomyocyte death (Rossi and Souza, 1999).

According to the experimental data, the best computational parameters of our model to describe the evolution of Chagas' disease were selected. They are the initial fraction of lattice-site occupation equaling 0.7, initial fraction of inflammatory cells equaling 2 x 10^{-3}, initial fraction of fibroblast equaling 2 x 10^{-3} and the number of inflammatory cells required for death of an inflammatory cell is greater than 3. The numerical value of inflammatory cells and fibroblast was

not compared to experimental data, because the value of these parameters was not determined in the experiment developed by Soares et al. (2001b).

4.4. Conclusion

In this article, we have presented a three-dimensional multi-agent-based model to represent the evolution of Chagas' disease. This model reproduces the acute and chronic phases of *T. cruzi* infection. Therefore, it simulates the kinetics of parasite nests, inflammatory cells, fibrosis progression and cardiomyocyte death. Also, it reproduces the volume occupied by all different types of cells in the myocardial tissue. The main aim is to understand the participation of different types of cells in the development of CChC.

Our results were compared with experimental data (Soares et al., 2001b), excepting for the kinetics of fibrosis progression and cardiomyocytes death, giving a good agreement. This implies that our model rules can be used to understand the evolution of Chagas' disease. Furthermore, different authors propose several hypotheses that have never been investigated by using a proper computational model. The results show that the microscopic rule of the autoimmune response is important for the modelling of CChC. Our data show that the initial fraction of lattice-site occupation and the number of inflammatory cells required for death of an inflammatory cell modify the kinetics of inflammatory cells and parasite nests. Also, the initial fraction of inflammatory cells has less influence on the kinetics of inflammatory cells and parasite nests. This occurs because the immunological system has a rapid autoregulatory response due to the migration [rule (i)] and death [rules (ii) and (iv)]. Finally, the initial fraction of fibroblasts modifies the fibrosis progression and the cardiomyocyte death.

Acknowledgements

This work was supported by FAPESB.

References

Andrade, L.O., Andrews, N.W. (2005) The *Trypanosoma cruzi*–host-cell interplay: location, invasion, retention. Nat. Rev. Microb., 3, 819–823.

Andrade, Z.A. (1999) Immunopathology of Chagas Disease. Mem. Inst. Oswaldo Cruz, 94, 71-80.

Coura, J.R. (2007) Chagas disease: what is known and what is needed – A background article. Mem. Inst. Oswaldo Cruz, 102, 113-122.

de Souza, E.M. et al. (2003) Host and parasite apoptosis following *Trypanosoma cruzi* infection in in vitro and in vivo models. Cell Tissue Res., 314, 223–235.

Galvão, V., Miranda, J.G.V., Ribeiro-dos-Santos, R. (2008) Development of a two-dimensional agent-based model for chronic chagasic cardiomyopathy after stem cell transplantation. Bioinformatics 24, 2051 – 2056.

Galvão, V., Miranda, J.G.V. (2009) Modeling the Chagas' disease after stem cell transplantation. Physica A, 388, 1747-1754.

Grimm, V. *et al.* (2006) A standard protocol for describing individual-based and agent-based models. Ecol. Model., 198, 115 - 126.

Higuchi, M.L. et al. (2003) Pathophysiology of the heart in Chagas' disease: current status and new developments. Cardiovasc. Res., 60, 96–107.

Isasi, S.C., Sibona, G.J., Condat, C.A. (2001) A simple model for the interaction between *T. cruzi* and its antibodies during Chagas infection. J. Theor. Biol., 208, 1–13.

Kelly, J.M. (2000) A B-cell activator in Chagas disease. Nat. Med., 6, 865–866.

Nelson, P., Velasco-Hernández, J.X. (2002) Modeling the immune response to parasitic infections: Leishmaniasis and Chagas disease. Comments Theor. Biol., 6, 161.

Rossi, M.A., Souza, A.C. (1999) Is apoptosis a mechanism of cell death of cardiomyocytes in chronic chagasic myocarditis? Int. J. Cardiol., 68, 325–331.

Sibona, G.J., Condat, C.A. (2002) Dynamic analysis of a parasite population model. Phys. Rev. E, 65, 031918.

Sibona, G.J., Condat, C.A., Isasi, S.C. (2005) Dynamics of the antibody-*T. cruzi* competition during Chagas infection: prognostic relevance of intracellular replication. Phys. Rev. E, 71, 020901.

Soares, M.B.P., Pontes-de-Carvalho, L., Ribeiro-dos-Santos, R. (2001a) The pathogenesis of Chagas' disease: when autoimmune and parasite-specific immune responses meet. An. Acad. Bras. Cienc., 73, 547-559.

Soares, M.B.P. et al. (2001b) Modulation of chagasic cardiomyopathy by interleukin-4: dissociation between inflammation and tissue parasitism. Am. J. Pathol., 159, 703–709.

Soares, M.P.B., et al. (2004) Transplanted bone marrow cells repair heart tissue and reduce myocarditis in chronic chagasic mice. Am. J. Pathol., 164, 441–447.

Teixeira, A.R.L, Nascimento, R.J., Sturm, N.R. (2006) Evolution and pathology in Chagas disease - A Review. Mem. Inst. Oswaldo Cruz, 101, 463-491.

Vinnakota, K.C., Bassingthwaighte, J.B. (2004) Myocardial density and composition: a basis for calculating intracellular metabolite concentrations. Am. J. Physiol. Heart Circ. Physiol., 286, H1742–H1749.

WHO-World Health Organization (2002) Control of Chagas' disease: Second report of a WHO Expert Committee. WHO Tech. Rep. Ser., 905, 1-109.

Wolfram, S. (1986) Theory and applications of cellular automata. World Scientific, Singapore.

CHAPTER 5

Modularity map of the network of human cell differentiation

Abstract

Cell differentiation in multicellular organisms is a complex process whose mechanism can be understood by a reductionist approach in which the individual processes that control the generation of different cell types are identified (Alberts et al., 2002; Freitas, 1999; Kirschstein and Skirboll, 2001; Sadler, 2004; Sell, 2004; Valentine, 2003; Vickaryous and Hall, 2006). Alternatively, a large scale approach in search of different organizational features of the growth stages promises to reveal the modular global structure (Caldarelli and Vespignani, 2007; Guimerà et al., 2005; Ravasz et al., 2002; Sadler, 2004; Valentine, 2003; Vickaryous and Hall, 2006) underlying the differentiation process and to therefore discover previously unknown relations between cell types. Here we sort and analyze a large set of scattered data to construct the network of human cell differentiation (NHCD) based on cell types (nodes) (Alberts et al., 2002; Valentine, 2003; Vickaryous and Hall, 2006) and differentiation steps (links) from the fertilized egg to a developed human. We discover a dynamical law of critical branching process which in turn reveals a surprinsing selfsimilar regularity in the modular organization of the network (Caldarelli and Vespignani, 2007; Guimerà et al., 2005; Ravasz et al., 2002; Song et al., 2005). Using this law we observe the network at different scales and identify hitherto unknown clusters of cell types (modules) following a hierarchical organization of specialized tissues and organs. Such a mapping would allow one to treat the development of a particular function in the context of the complex network of human development as a whole. Our results point to an integrated large-scale view of the dynamical and topological properties of the network of human cell differentiation disclosing important features of the modular organization during the embryonic development.

5.1. Manuscript text

The cell differentiation process plays a crucial role in the embryonic development of multicellular organisms. Recent advances in the research on stem cell properties and embryonic

development have uncovered several steps in the differentiation process (Alberts et al., 2002; Freitas, 1999; Kirschstein and Skirboll, 2001; Sadler, 2004; Sell, 2004). Single and multiple sequences of cell differentiation have been identified through in-vivo observations of a particular embryo during early stages of development and through pathology studies of miscarriages during late stages of the process. While the identification of each cell differentiation step has been the subject of intense research, an integrated view of this complex process is still missing. Such a global view promises to reveal features associated with the large-scale modular organization of the cell types with the purpose of discovering new functional modules between cell types using novel theoretical network analysis. In this letter, we take advantage of the current knowledge on the sequence of cell differentiation processes, which is spread over a vast specialized literature (see Table I and references therein), to reveal and characterize the topological and dynamical features associated with the network of human cell differentiation (NHCD).

We construct the NHCD by systematically gathering the scattered information on the evolution of each cell type present in the embryo from a predecessor with a higher degree of differentiation potential into a more specialized type. The process of cell differentiation is then mapped onto a complex network which consists of 873 nodes connected through 977 edges. The nodes in the network represent distinct cell types reported in the literature and the edges represent the association between two cell types through a differentiation event. The initial steps of the NHCD are shown in Fig. 1, while the resulting network structure is shown in Fig. 2 and Figs 6a and 6b. Each cell type appears as one node and bifurcation points give rise to the connections in the network. Certain types of cells can be generated following more than one path from the fertilized egg, generating some closed loops of edges in the network. The NHCD comprises 529 branches of different sizes with each branch ending in a tree leaf, indicating cell types that do not undergo further differentiation. The surviving cell types are denoted by filled circles, while non-surviving ones are indicated by empty circles. Note however, that the most recent compilation of cell types done in (Vickaryous and Hall, 2006) reports only 407 cell types in normal, healthy, human adults (surviving cell types). Therefore, not all branch endpoints correspond to cell types in born humans. The complete collected data is listed in the Supplementary Information, with an enumeration of cells and links between the cell types, their time of appearance in days after fecundation (T_a), and the reference to the publications reporting each link. To the best of our knowledge, the structure identified here provides the most complete schematic diagram of the human differentiation process to date.

It is visually apparent from Fig. 1b that the NHCD has a prominent modular structure. The continuous differentiation of cells into more specialized functions naturally leads to the formation of dense isolated clusters in the NHCD. As a first approach to understand this modular structure we

cluster cell types in the network of Fig. 2 according to their known biological functions; different colors indicate 19 functional modules (C1 − C19) (Sadler, 2004) (see Table I and references therein). The largest communities were extracted from Refs. (Alberts et al., 2002; Anglani et al., 2004; Bianco et al., 2001; Chen and Goldhamer, 2003; Freitas, 1999; Forge and Wright, 2002; Herrick and Mutsaers, 2004; Horster et al., 1999; Janeway et al., 2001; Jessen and Mirsky, 2005; Kirschstein and Skirboll, 2001; Nakashima and Redid, 2003; Otto, 2002; Panteleyev et al., 2001; Paxinos and Mai, 2004; Sadler, 2004; Santagati and Rijli, 2003; Savage et al., 2003; Sell, 2004; Temple, 2001; Vickaryous and Hall, 2006). There is, however, a certain degree of arbitrariness in this modular structure as the separation of the nodes into communities in our dataset is not unique. For instance, community C12, the neural lineage, could be divided into two sub-communities, representing the neural and the supporting (glial) cells (Freitas, 1999; Kirschstein and Skirboll, 2001; Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Temple, 2001; Vickaryous and Hall, 2006). On the other hand, the neural system module could be merged with the eye system module (Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006) on a larger scale. Therefore, a finer or coarser community structure can be extracted from the data when we look at the whole network at different scales of observation; an appropriate algorithm is needed to identify these communities in a systematic way.

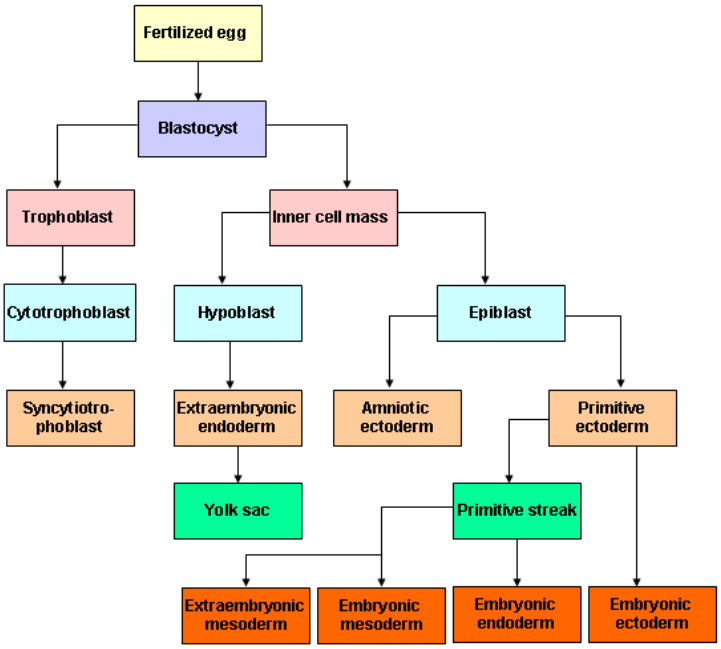


Fig. 1. The first steps of the NHCD construction. These steps, known to also be present in the formation of the majority of multicellular organisms, include the first cleavage of a fertilized egg, which is subsequently followed by the ball stage and the formation of primary germ cell layers, namely, the ectoderm, mesoderm, and endoderm. The fertilized egg is a totipotent stem cell and the predecessor of the blastocyst is a pluripotent stem cell. The blastocyst, in turn, gives rise to both trophoblast and inner cell mass. These two cells further differentiate into other types of cells, and so on.

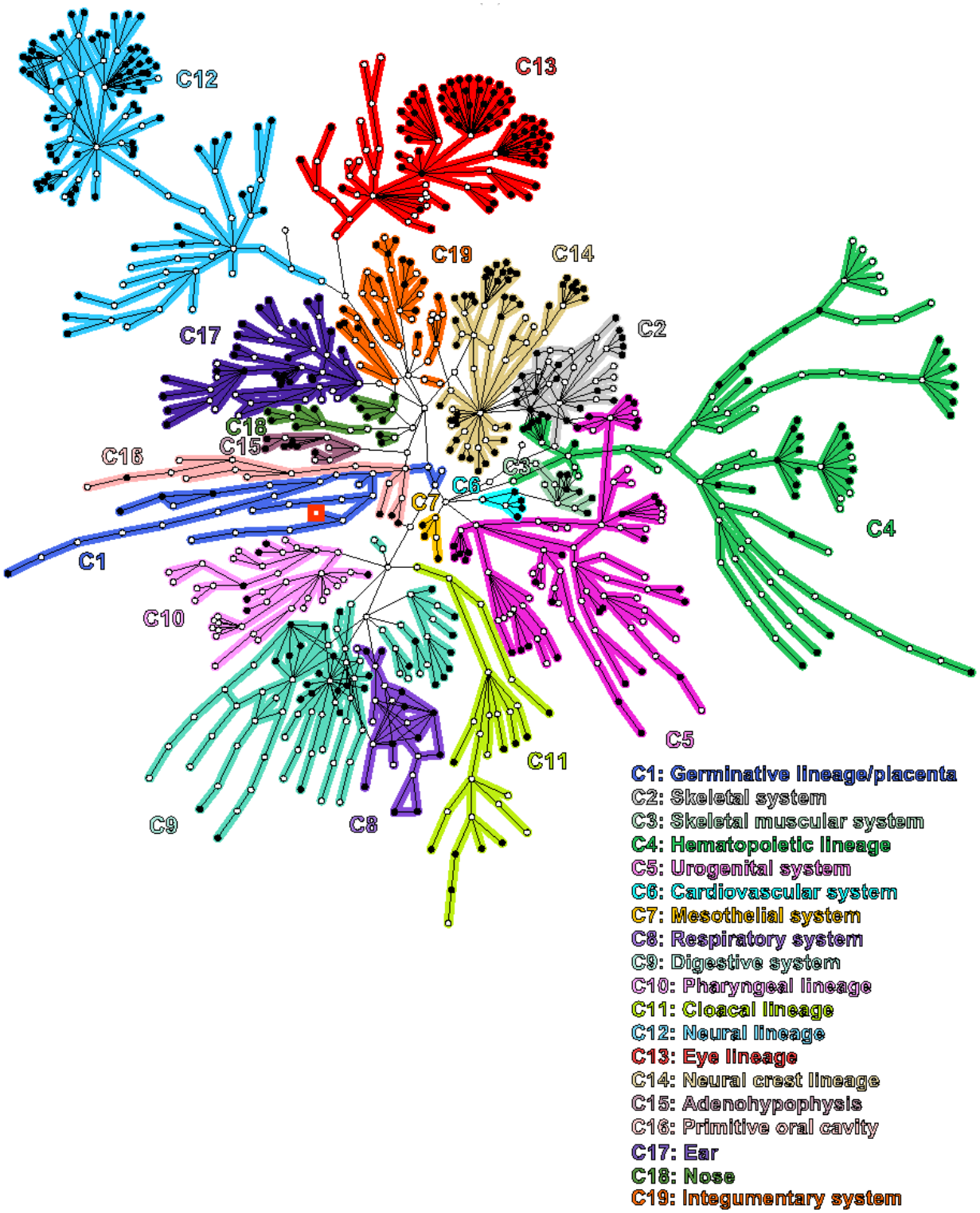


Fig. 2. Complex network representation of the human cell differentiation process. Following the above process until the fetus is fully developed yields the complex network shown in this figure. Each node, plotted as a circle, corresponds to a cell type and the edges to a differentiation step. The entire network originates from the fertilized egg (denoted by a red square) and leads to the specialized cells of a developed human. The links emerging from each cell type follow a bifurcation pattern, including a small number of edges that allow for the presence of loops. Filled circles correspond to nodes that survive at the end of the development process, while empty circles correspond to non-surviving cell types. Communities with distinct biological functions are indicated by different colors (see Table I for association of the branch labels to biological function C1 to C19).

Graph theoretical concepts allow us to unravel the scale dependence of the modular structure of the NHCD. Graph theory (Caldarelli and Vespignani, 2007) defines the length between two nodes (also called the chemical distance) as the number of links along the shortest path between the nodes in the network. We use this notion to identify modules of maximum length ℓ composed of highly connected cell types. This is done by assigning each node to a module or box after tiling the network with the smallest possible number of modules, N_B of size ℓ (all nodes in a module are at distance smaller than ℓ) (Song et al., 2005). This process results in an (NP-complete) optimization problem which can be solved by the use of the box-covering algorithm reported in (Song et al., 2007). We have also applied other community-detection algorithms such as the successive elimination of edges with highest betweenness centrality (Caldarelli and Vespignani, 2007; Newman and Girvan, 2004) and the resulting modular structure is similar to the one displayed in Fig. 2

The requirement of minimal number of modules to cover the network (N_B) guarantees that the partition of the network is such that each module contains the largest possible number of nodes and links inside the module with the constraint that the modules cannot exceed size ℓ. This optimized tiling process gives rise to modules with the fewest number of links connecting to other modules implying that the degree of modularity, defined by (Caldarelli and Vespignani, 2007; Gallos et al., 2007; Girvan and Newman, 2002; Guimerà and Amaral, 2005)

$$M(\ell) \equiv \frac{1}{N_B} \sum_{i=1}^{N_B} \frac{L_i^{in}}{L_i^{out}} \tag{1}$$

is maximized. Here L_i^{in} and L_i^{out} represent the number of links that start in a given module i and end either within or outside i, respectively. Large values of $M(L_i^{out} \to 0)$ correspond to a higher degree of modularity. The value of the modularity of the network M varies with ℓ, so that we can detect the dependence of modularity on different length scales, or equivalently how the modules themselves are organized into larger modules that enhance the degree of modularity.

For a given ℓ, we obtain the optimal coverage of the network with N_B modules (the full list of modules is in the SI). Analysis of the modularity Eq. (1) of the NHCD approximately yields a power-law functional form:

$$M(\ell) \sim \ell^{d_M} \tag{2}$$

which is detected through the modularity exponent d_M. We characterize the network using different snapshots in time and we find that $d_M \approx 2.0$ is approximately constant over the time evolution (Fig. 3a. This value reveals a considerable degree of modularity in the system (for comparison, a random network has $d_M = 0$ and a uniform lattice has $d_M = 1$ (Gallos et al., 2007)), as already evidenced by the network structure in Fig. 2. The lack of a characteristic length evidenced in the power law form of Eq. (2) suggests that the modules appear at all length-scales, i.e. modules are organized within larger modules in a self-similar way, so that the inter-connections between those clusters do not modify the basic modular character of the entire NHCD.

Thus, the NHCD presents a self-similar topology and remains statistically invariant when observed at different scales. Varying the module size ℓ yields the scaling relation for the number of modules (Fig. 3b):

$$N_B(\ell) \sim \ell^{-d_B} \tag{3}$$

where d_B represents the fractal dimension of the network (Song et al., 2005). We find that the fractal character is established at the early stages, yielding $d_B \approx 1.4$ as early as 30 days (Fig. 3b). As the network evolves, the fractal dimension increases slightly and finally reaches $d_B \approx 1.9$. The change in d_B implies that the complexity of the network increases, while it also approaches the form of a critical percolation network (Caldarelli and Vespignani, 2007), which is known to have $d_B = 2.0$.

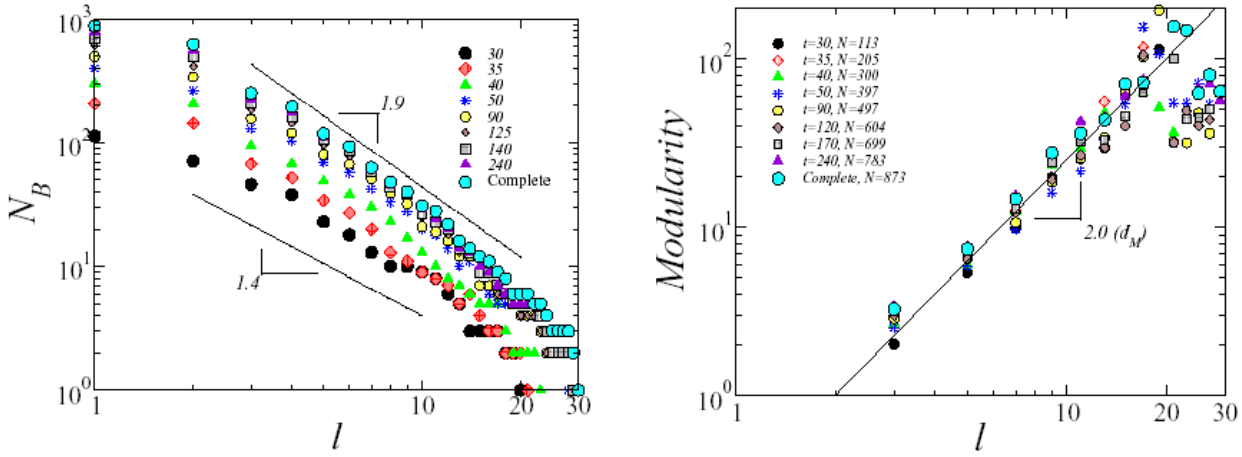


Fig. 3. Modular properties of the NHCD. a) Degree of modularity of the network, $M(\ell)$ at different times T_a (indicated in the figure) as a function of the scale of observation, ℓ. b) Number of boxes/modules, N_B, of maximum length ℓ, as identified by the box-covering algorithm (see MEMB, Maximum-Excluded-Mass-Burning, algorithm in (Song et al., 2007)) needed to cover the network as a function of ` for di®erent networks at time T_a.

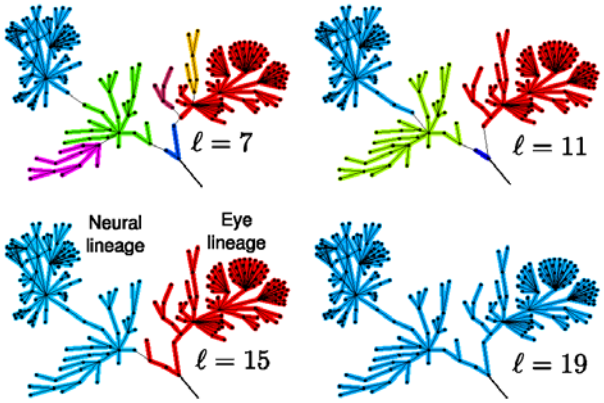
The significance of Eqs. (2) and (3) is that the modules need to be interpreted at a given length-scale. Figure 4 shows an example of such organization (Fig. 4b and Fig. 7 show the full modular structure). The entire eye lineage (Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006) is detected by the box-covering algorithm at $\ell = 11$, while the neural lineage appears as a single module at $\ell = 15$ (Freitas, 1999; Kirschstein and Skirboll, 2001; Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Temple, 2001; Vickaryous and Hall, 2006). Coarser and finer novel modules are identified by the algorithm. For instance, at $\ell = 11$ the neural lineage is split into the main neural and the supporting glial cell modules. When we increase the length to $\ell = 19$, the eye and neural system form a single meta-module while for $\ell = 7$ sub-modules are identified as shown in Fig. 4a. Thus, each cell type is connected to other types under a single principle according to which groups of nodes of all sizes self-organize according to one universal dynamics. This property allows us to renormalize the network by replacing each module by a single supernode to identify the network of modules as shown in Fig. 4b.

The dynamics leading to such a structure can be unraveled by the study of the NHCD as a growth process. The knowledge of the time of appearance of each cell type, T_a, makes it possible to follow the cumulative growth of the embryo in terms of the number of cell types that appear during the growth, $N(t)$, as well as the number of cell types that eventually survive in the organism (Fig. 5a). As expected, surviving cells emerge in the later stages of the gestation period (Fig. 5a). However, the difference between the total and the surviving number of cell types indicates that generation of new types of non-surviving cells takes place even during the final gestation months.

The increase of the network size, $N(t)$, is initially exponential and after $t_* = 40$ days changes into a slower growth (Fig. 5a). Only a small percentage of the nodes grow within a given time interval, so that the network activity is focused in a small number of them at a given time. The number of nodes that differentiate at a given time are shown in Fig. 5b. We observe an activity that increases monotonically up to around $t_* = 40$ days and then drops to lower values. The cross-over time $t_* = 40$ days observed in Figs. 5a and 5b separates two regimes of growth and corresponds to the time below which most of the cells have a plastic characteristic (i.e., the capability to differentiate) and above which they become functional (Sell, 2004). Interestingly, the two regimes observed in $N(t)$ merge into a single universal functional curve when we replot $N(\ell_{N1})$ as a function of the chemical distance to the fertilized egg, ℓ_{N1} (Fig. 5c). This result suggests that the topological distance in the network ℓ_{N1} is the natural variable to characterize the growth process in a universal form rather than the time. The dynamic of $N(\ell)$ follows a typical logistic (Verhulst) process of population growth where the rate of growth is restricted by environmental limitations:

$dN/d\ell = rN[1-N/N_f]$, with solution, $N(\ell) = N_f \exp(r\ell)/(N_f + (\exp(r\ell)-1))$ (see the fitting in Fig. 5c) where N_f is the final number of cell types and $r = 0.65$ is the growth rate of cell types.

a)



b)

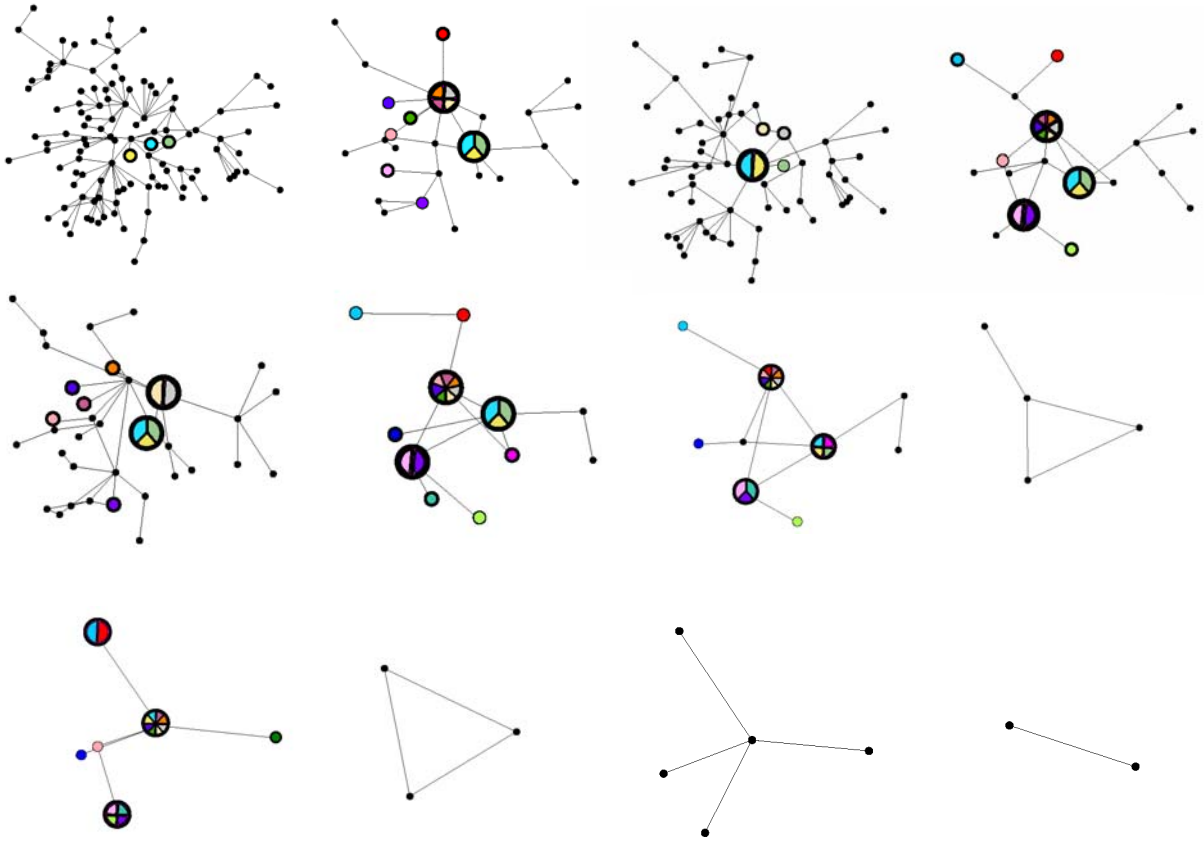


Fig. 4. Detection of modules at different scales. a) Modules are detected by the box-covering algorithm (MEMB in (Song et al., 2007)) to tile the network with NB boxes of size ℓ. We show the identified modules corresponding to the functional classes C12 : neural system and C13: eye system (full modular structure is in Fig. 4b and Fig. 7). Both functional classes appear at $\ell = 15$. At other scales the box-covering algorithm detects new functional relations between cell types expressed in the obtained meta-modules and sub-modules. For instance, at $\ell = 11$ the neural lineage is further divided into two modules. The obtained scale-dependent families of nodes satisfy the scaling laws Eqs. (2) and (3). b) The network of modules at different ℓ values. Small circles with thick lines indicate the smallest ℓ value for which a functional module is detected through MEMB. Small circles with thin lines show that a MEMB box corresponds to a specific function at a larger ℓ value. Two or more functional modules that are joined into a larger MEMB box are presented through a pie chart with all the corresponding colors.

Analysis of the network connectivity reveals that the average number of links per node in the final stages of the entire NHCD is $\langle k \rangle = 2.24$ (Fig. 5d). Even though $\langle k \rangle \approx 2$, there is a broad degree distribution (scale-free (Caldarelli and Vespignani, 2007), $P(k) \sim k^{-\gamma}$, $\gamma \approx 2.75$, Fig. 8). This implies that there is always a small number of crucial cell types that differentiate much more than the others, a fact that agrees with evidence on the existence of a few cells with large plasticity potential. As this potential is rapidly lost after 40 days, cell types change their development ability in favor of the organism life maintenance.

The fact that the average degree is close to 2 implies that the dynamical evolution of NHCD can be described by a critical branching process where every node has a certain probability of generating offsprings, in which case the critical condition for the branching to continue is $\langle k \rangle = 2$. This effectively means that each node needs to give at least onedescendant in order for the network to keep growing. If $\langle k \rangle < 2$, the growth would stop early, while for $\langle k \rangle > 2$ the growth would be faster than exponential.

The network reaches the condition of criticality, $\langle k \rangle \approx 2$, at around $t_* = 40$ days (Fig. 5d) in conjunction with the transition from plasticity to functional behavior. After this, the average degree remains just above criticality to sustain a growth rate that guarantees the network survival. The majority of the nodes propagate the growth in a single line, but there are nodes which generate significantly more descendants to generate the diversity implied by the power-law distributions of degree and modularity. In summary, we find that the interplay between cell plasticity and specialization leads to a complex network of cell types and differentiation steps that self-organize into a modular structure. This particular architecture isolates cell types in well-defined modules and evolves following the stable laws of a critical branching process. The self-similar modular features evidence in Fig. 3 is established early in the process and remain invariant during the evolution of the NHCD, although the network size changes significantly. This information could be of importance in providing predictions of functional attributes to a number of identified modules of cell types in the NHCD.

Fig. 6 provides an alternate representation of the NHCD. Different cell types appearing as bifurcation points of the network are stacked along the vertical axis. In Fig. 6a the horizontal axis corresponds to the shortest path ℓ_{N1} calculated from node $N1$, the fertilized egg, to any given node, while in Fig. 6b the same network is shown as a function of the node appearance time T_a. The links emerging from each cell type follow a bifurcation pattern ending up at the right side with $k-1$ branches, each one of them representing one of the more specialized cell types. Red nodes correspond to the surviving cell types, and they obviously appear at later times, while the non-

surviving cell types, the blue nodes, emerge during the early stages of the process. The color of the edges corresponds to one of the 19 function groups identified in Table I. Links that generate loops are plotted in red. Figure 6b contains the same information as Fig. 6a but we plot each cell type according to its time of appearance rather than as a function of the chemical distance to $N1$, as in Fig. 6a. The white and yellow alternating vertical stripes divide the time axis in intervals in days. The branches have been extended so that each cell appears only in the corresponding interval. Colors and labels are the same as in Fig. 6a.

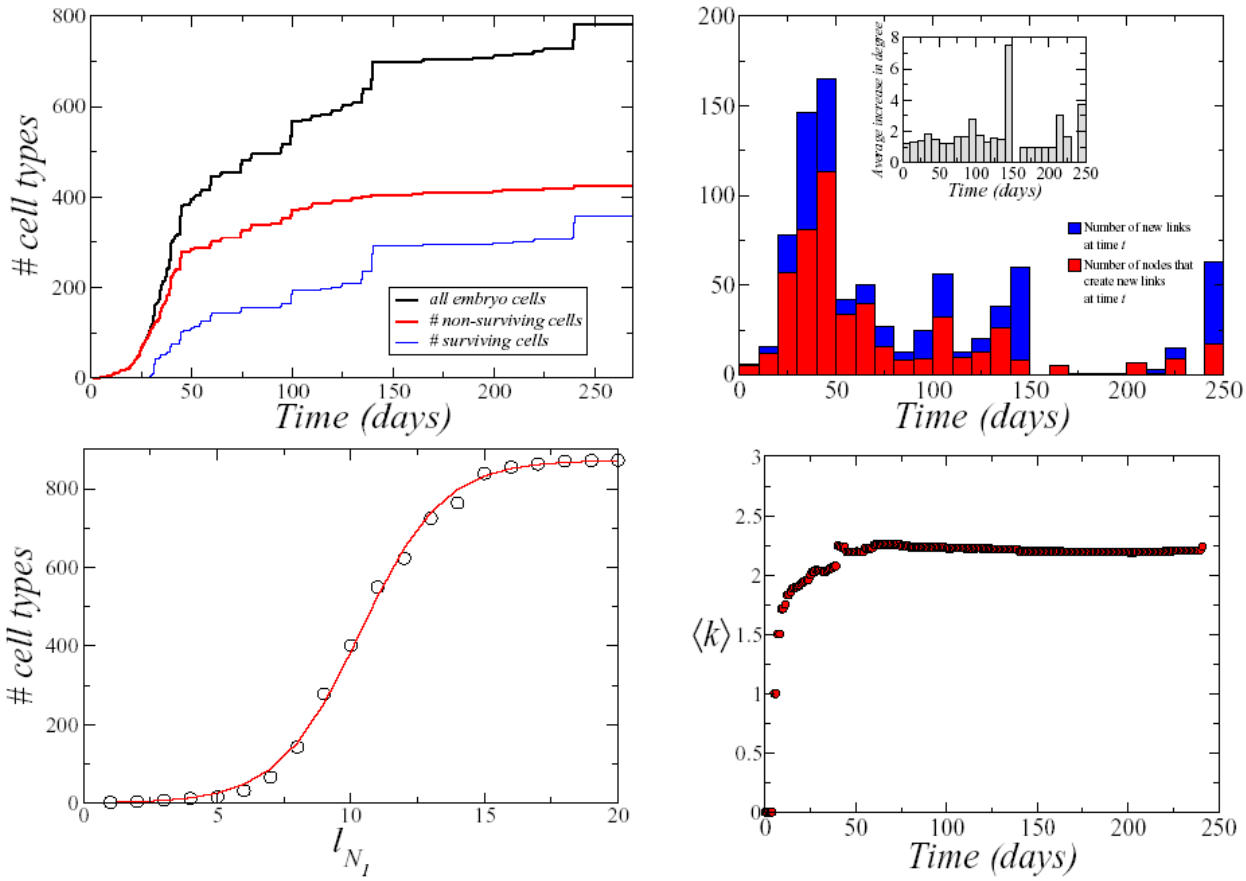
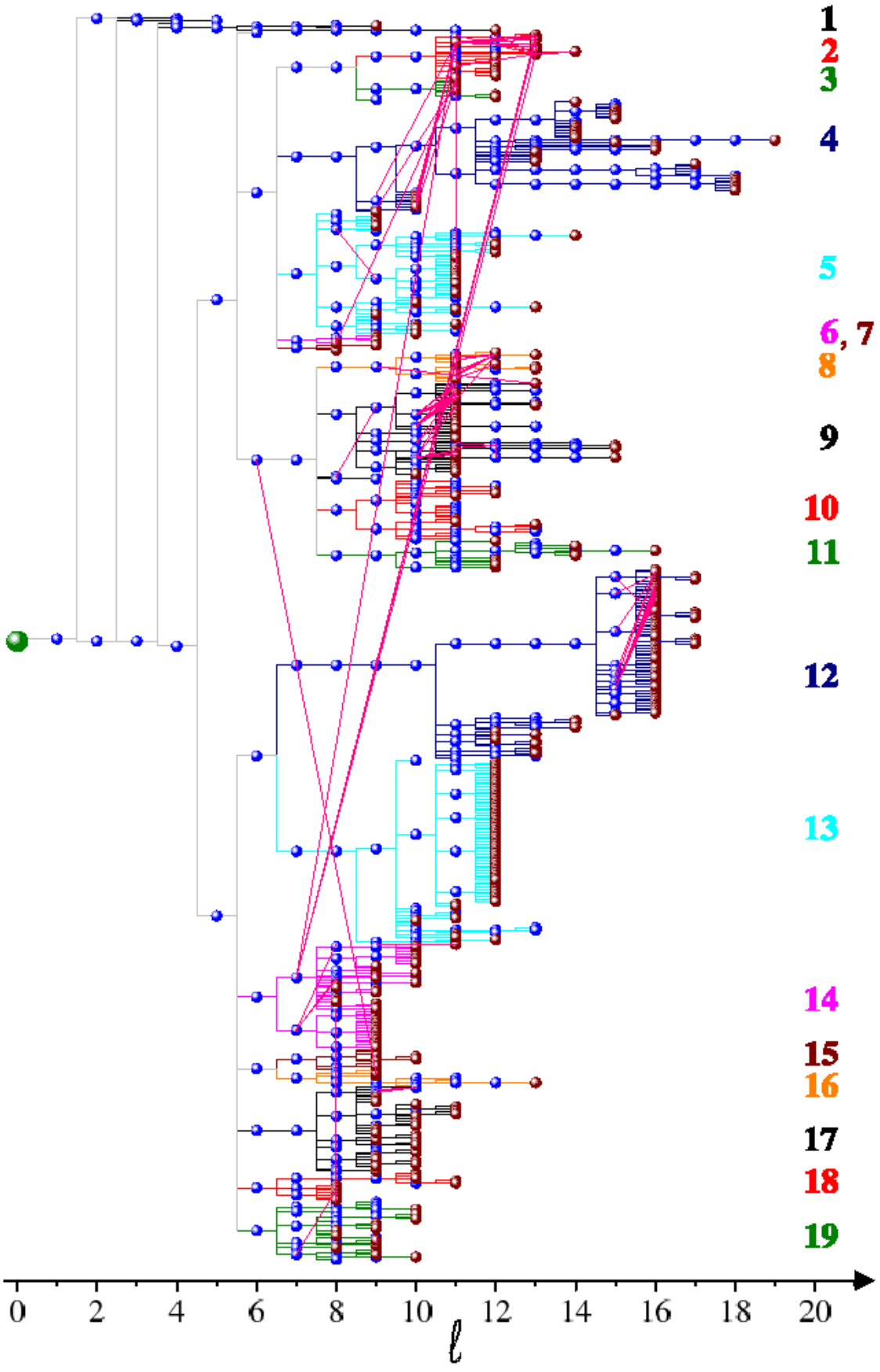


Fig. 5. Growth properties of the NHCD. a) Number of cell types in the network, N(t), as a function of time. We were able to find precise information about the appearance time T_a for 782 among the 873 cell types. Those cells with missing appearance time have not been taken into account in this plot. Also shown are the time evolution of the number of surviving and non-surviving cells. b) Number of nodes whose degree increases at time t (red histogram) and number of new links appearing in the network (blue histogram) as a function of time. If all nodes were giving just one child then the two histograms would coincide. Inset: The average number of new links per node at a given time can be found by dividing the two histograms in the main plot. This plot shows how intense is the activity at that particular time. Despite the variation in activity, the new connections average around 1, which gives a critical branching ratio of $\langle k \rangle \approx 2$. c) Number of cell types versus the chemical distance to the first node, ℓ_{N1}. This distance is only determined by the connections between the cell types, and is not influenced by the appearance time, so that we include all 873 cell types. d) Average degree $\langle k \rangle$ in the network as a function of time showing that the network achieves the condition of critical branching process $\langle k \rangle \approx 2$ at around $t_* = 40$.

a)



b)

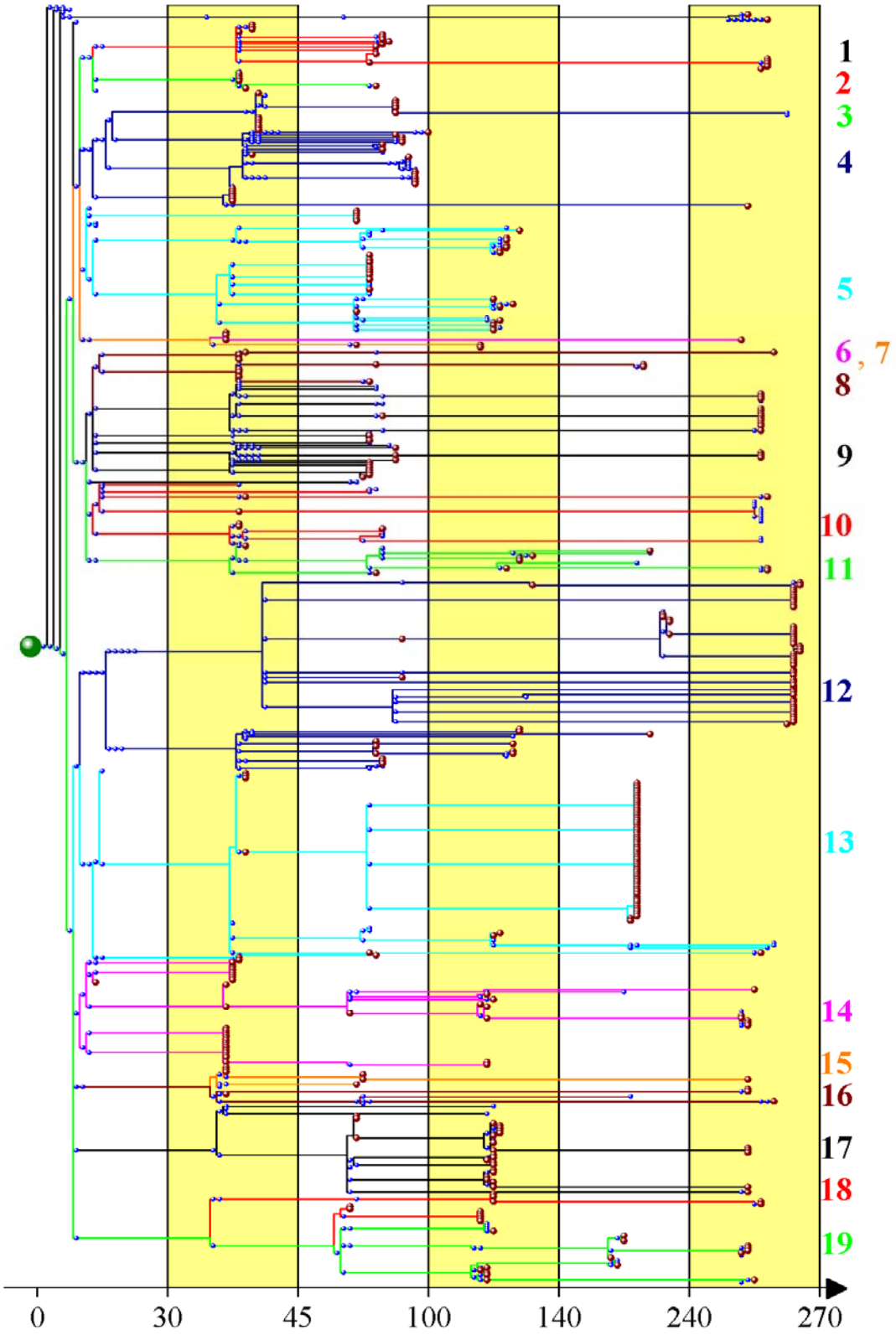


Fig. 6. Alternative representation of the NHCD. a) The horizontal axis measures the shortest path from each cell type to the fertilized egg along the network, ℓ_{N1} . b) The horizontal axis denotes the appearance time of a given cell type, T_a .

The fertilized egg is followed by the ball stage, and the formation of the primary germ cell layers. Currently, it is known that until the ball stage, cell division is symmetric and produces further totipotent stem cells (Sell, 2004). These cells then give rise to all the differentiated tissues of the organism as well as the extra-embryonic tissues (placenta, umbilical cord, etc.). Moreover, in the course of the entire process of organism formation, there is a monotonic decrease in the differentiation potential ($totipotent \rightarrow pluripotent \rightarrow multipotent \rightarrow unipotent$) accompanied with an increase in cell specialization. The number of distinct cell types in a human also depends on age and gender. Not all 873 cell types are present in a human being. Among those absent are the placenta cells that are generated from the fertilized egg during embryo development, as well as other somatic cell types that are important to control embryo and fetus development. The catalog presented in Ref. (Vickaryous and Hall, 2006) reports 407 distinct cell types in the human body, all of which can be identified in the network representation. Most of them occupy the end points of the 529 branches in Figs. 2 and 6. Therefore, not all tree leaves (branch endpoints) correspond to cell types in born humans. The average shortest path calculated from all cell types to $N1$ is $\langle \ell_{N1} \rangle = 10.93$, as expected from the large concentration of links in the interval $8 \leq \ell \leq 13$ (Fig. 6a). The network representation of Fig. 6 should not be interpreted as a typical hierarchical tree where only the leaves represent the real nodes and the branches indicate community clustering. It is rather an evolutionary network indicating the differentiation of a cell type into another, where ancestor nodes do not necessarily disappear and they can still participate in the subsequent evolution of the network. Although the network is represented like a dendrogram, it should not be read as such geometrical construct. Dendrograms usually indicate community splitting, starting at the left side of a plot where a single community contains all nodes.

Figure 7 shows the full modular structure of the NHCD as detected by the boxcovering algorithm at different length scales. A detail of this process is represented in Fig. 4. Figure 8 shows the degree distribution of the NHCD for different times. Table I lists the different functional modules of the NHCD and the respective citations to the literature. The complete collected data is listed in Supplementary information. Table II includes all the cell types and the Table III includes the links between the cell types their time of appearance in days after fecundation (T_a), and the reference to the publications reporting each link. Data on the structure of individual communities were obtained from the specialized literature. The most important ones contain data from several communities, like germ layer and extraembryonic tissue (Alberts et al., 2002; Kirschstein and Skirboll, 2001; Sadler, 2004; Sell, 2004), digestive system (Sadler, 2004; Sell, 2004), pharyngeal system (Freitas, 1999; Sadler, 2004; Vickaryous and Hall, 2006), cloacal system (Sadler, 2004), neural system (Sell, 2004; Vickaryous and Hall, 2006), eye (Sadler, 2004; Vickaryous and Hall,

2006), primitive oral cavity (Sadler, 2004), nose (Sadler, 2004), skeletal system (Bianco et al., 2001; Freitas, 1999), skeletal muscular system (Chen and Goldhamer, 2003; Sell, 2004; Vickaryous and Hall, 2006), hematopoietic system (Janeway et al., 2001), urogenital system (Anglani et al., 2004; Horster et al., 1999), cardiovascular system (Sadler, 2004; Vickaryous and Hall, 2006), mesothelium (Herrick and Mutsaers, 2004), respiratory (Otto, 2002), neural crest (Jessen and Mirsky, 2005; Nakashima and Redid, 2003; Sadler, 2004; Santagati and Rijli, 2003), adenohypophysis (Savage et al., 2003), ear (Forge and Wright, 2002; Vickaryous and Hall, 2006), and integumentary system (Panteleyev et al., 2001; Sadler, 2004; Vickaryous and Hall, 2006).

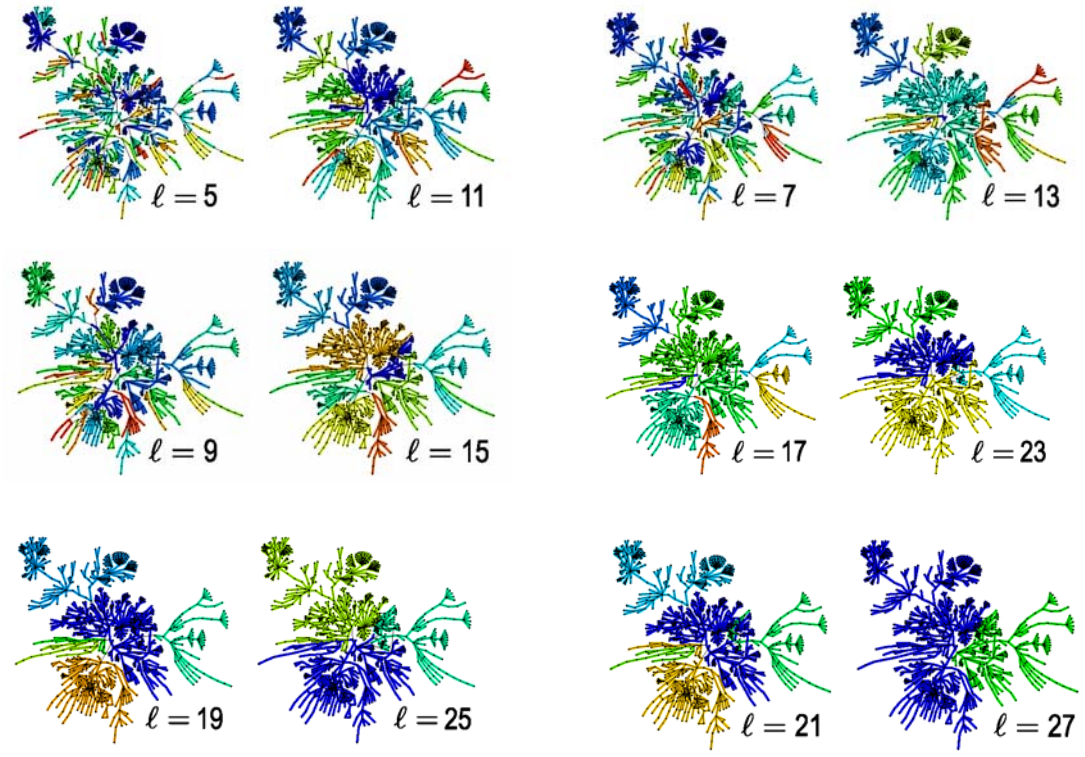


Fig. 7. Full modular structure of the NHCD at the indicated length ℓ. Each node is depicted with a different color indicating the module to which it belongs to.

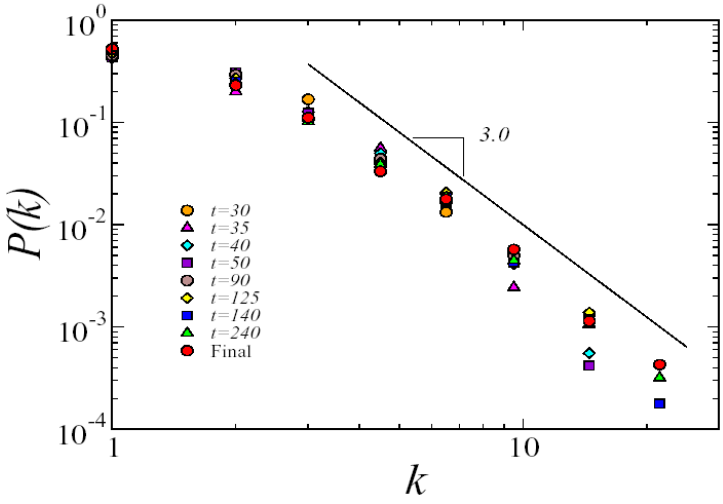


Fig. 8. Degree distribution $P(k)$ for the NHCD at different growth stages, from $t = 30$ days to $t = 240$ days.

TABLE I. Identification of the distinct biological functions of the network branches indicated in Figs. 2 and 3. The third column lists the used references for building the NHCD.

Label	Biological function	Reference
1	Germinative Lineage and Placenta	Alberts et al., 2002; Kirschstein and Skirboll, 2001; Sadler, 2004; Sell, 2004
2	Skeletal System	Bianco et al., 2001; Freitas, 1999; Mochida, 2005; Sadler, 2004; Sell, 2004; Sanders et al., 1999; Sell, 2004; Towler and Gelberman, 2006 ; Vickaryous and Hall, 2006
3	Skeletal Muscular System	Chen and Goldhamer, 2002; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
4	Hematopoietic Lineage	Alberts et al., 2002; Kirschstein and Skirboll, 2001; Janeway et al., 2001; Minasi et al., 2002; Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
5	Urogenital System	Anglani et al., 2004; Coulter, 2004; Horster et al., 1999; Lopez et al., 2001; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
6	Cardiovascular System	Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
7	Mesothelial Lineage	Herrick and Mutsaers, 2004; Sadler, 2004
8	Respiratory System	Freitas, 1999; Otto, 2002; Sadler, 2004; Sell, 2004
9	Digestive System	Bardeesy and DePinho, 2002; Fausto, 2004; Freitas, 1999; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
10	Pharyngeal Lineage	Blackburn and Manley, 2004; Freitas, 1999; Sadler, 2004; Vickaryous and Hall, 2006
11	Cloacal Lineage	Foster et al., 2002; Freitas, 1999; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
12	Neural Lineage	Freitas, 1999; Kirschstein and Skirboll, 2001; Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Temple, 2001; Vickaryous and Hall, 2006
13	Eye Lineage	Paxinos and Mai, 2004; Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
14	Neural Crest Lineage	Jessen and Mirsky, 2005; Nakashima and Redid, 2003; Sadler, 2004; Santagati and Rijli, 2003; Sell, Szeder et al., 2003; Vickaryous and Hall, 2006
15	Adenohypophysis	Paxinos and Mai, 2004; Sadler, 2004; Savage et al., 2003; Vickaryous and Hall, 2006
16	Primitive Oral Cavity	Freitas, 1999; Nakashima and Redid, 2003; Sadler, 2004; Vickaryous and Hall, 2006
17	Ear	Forge and Wright, 2002; Freitas, 1999; Paxinos and Mai, 2004; Sadler, 2004; Vickaryous and Hall, 2006
18	Nose	Freitas, 1999; Sadler, 2004; Vickaryous and Hall, 2006
19	Integumentary System	Freitas, 1999; Hennighausen and Robinson, 2005; Panteleyev et al., 2001; Potten and Booth, 2002; Sadler, 2004; Stoeckelhuber et al., 2003; Vickaryous and Hall, 2006

Acknowledgements

We thank J. Brujic, B. Dubin-Thaler, H.D. Rozenfeld, S. Havlin, and T. Rattei for valuable discussions. This work was supported by National Science Foundation Grant SES-0624116.

References

Alberts, B. et al. (2002) Molecular Biology of the Cell. Fourth ed., Garland Science, New York.

Anglani, F., Forino, M., Del Prete, D., Tosetto, E., Torregrossa, R., D'Angelo, A. (2004) In search of adult renal stem cells. J. Cell. Mol. Med., 8, 474-487.

Bardeesy, N., DePinho, R.A. (2002) Pancreatic cancer biology and genetics. Nat. Rev. Cancer, 2, 897-909.

Bianco, P., Riminucci, M., Gronthos, S., Robey, P.G. (2001) Bone marrow stromal stem cells: nature, biology, and potential applications. Stem Cells, 19, 180–192.

Blackburn, C.C., Manley, N.R. (2004) Developing a new paradigm for thymus organogenesis. Nat. Rev. Immunol., 4, 278-289.

Caldarelli G., Vespignani A. (eds). (2007) Large scale structure and dynamics of complex networks. World Scientific, Singapore.

Chen, J.C.J., Goldhamer, D.J. (2003) Skeletal muscle stem cells. Reprod. Biol. Endocrinol. 1, 101.

Coulter, C.L. (2004) Functional biology of the primate fetal adrenal gland: advances in technology provide new insight. Clin. Exp. Pharmacol. P., 31, 475-484.

Fausto, N. (2004) Liver regeneration and repair: hepatocytes, progenitor cells, and stem cells. Hepatology, 39, 1477-1487.

Freitas, R.A., Jr. (1999) Nanomedicine, Volume I: Basic Capabilities. Landes Bioscience, Georgetown, Texas.

Forge, A., Wright, T. (2002) The molecular architecture of the inner ear. Brit. Med. Bull., 63, 5–24.

Foster, C.S., Dodson, A., Karavana, V., Smith, P.H., Ke, Y. (2002) Prostatic stem cells. J. Pathol., 197, 551-565.

Gallos, L.K., Song, C., Havlin, S., Makse, H.A. (2007) Scaling theory of transport in biological complex networks. Proc. Nat. Acad. Sci. USA, 104, 7746–7751.

Girvan M., Newman, M.E.J. (2002) Community structure in social and biological networks. Proc. Nat. Acad. Sci. USA, 99, 7821–7826.

Guimerà, R., Amaral, L.A.N. (2005) Functional cartography of complex metabolic networks. Nature, 433, 895–900.

Hennighausen, L., Robinson, G.W. (2005) Information networks in the mammary gland. Nat. Rev. Mol. Cell Bio., 6, 715 – 725.

Herrick, S.E., Mutsaers, S.E. (2004) Mesothelial progenitor cells and their potential in tissue engineering. Int. J. Biochem. Cell B., 36, 621-642.

Horster, M.F., Braun, G.S., Huber, S.M. (1999) Embryonic renal epithelia: induction, nephrogenesis, and cell differentiation. Physiol. Rev., 79, 1157-1191.

Kirschstein, R., Skirboll, L. R. (2001) Stem cells: scientific progress and future research directions. NIH, Bethesda.

Janeway, C.A., Travers, P., Walport, M. & Shlomchik, M. (2001) Immunobiology: the immune system in health and disease. Fifth ed., Garland Science.

Jessen, K.R., Mirsky, R. (2005) The origin and development of glial cells in peripheral nerves. Nat. Rev. Neurosci., 6, 671–682.

Lopez, M.L.S.S, Pentz, E.S., Robert, B., Abrahamson, D.R., Gmez, R.A. (2001) Embryonic origin and lineage of juxtaglomerular cells. Am. J. Physiol. Renal Physiol. 281, 345-356.

Minasi, M.G., et al. The meso-angioblast: a multipotent, self-renewing cell that originates from the dorsal aorta and differentiates into most mesodermal tissues. Development, 129, 2773-2783 (2002).

Mochida, J. (2005) New strategies for disc repair: novel preclinical trials. J. Orthop. Sci., 10, 112–118.

Nakashima, M., Redid, A.H. (2003) The application of bone morphogenetic proteins to dental tissue ingineering. Nat. Biotechnol., 21, 1025–1032.

Newman, M.E.J., Girvan M. (2004) Finding and evaluating community structure in networks. Phys. Rev. E, 69, 026113.

Otto, W.R. (2002) Lung epithelial stem cells. J. Pathol., 197, 527–535.

Panteleyev, A., Jahoda, C.A.B., Christiano, A.M. (2001) Hair follicle predetermination. J. Cell. Sci., 114, 3419-3431.

Paxinos, G., Mai J.K. (2004) The Human Nervous System. Second ed., Elsevier Academic Press.

Potten, C.S., Booth, C. (2002) Keratinocyte Stem Cells: a Commentary. J. Invest. Detmatol., 119, 888-899.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N. & Barabási, A.-L. (2002) Hierarchical organization of modularity in metabolic networks. Science, 297, 1551-1555.

Sadler, T.W. (2004) Langman's Medical Embryology. Nineth ed., Lippincott Williams & Wilkins, Baltimore.

Santagati, F., Rijli, F.M. (2003) Cranial neural crest and the building of the vertebrate head. Nat. Rev. Neurosci., 4, 806–818.

Savage, J.J., Yaden, B.C. Kiratipranon, P., Rhodes, S.J. (2003) Transcriptional control during mammalian anterior pituitary development. Gene, 319, 1–19.

Sell, S. (2004) Stem Cells Handbook. Humana Press, Totowa, NJ.

Song, C., Gallos, L.K., Havlin, S., Makse, H.A. (2007) How to calculate the fractal dimension of a complex network. J. Stat. Mech: Theory and Experiments, P03006.

Song, C., Havlin, S., Makse, H.A. (2005) Self-similarity of complex networks. Nature, 433, 392–395.

Stoeckelhuber, M., Stoeckelhuber, B.M., Welsch, U. (2003) Human Glands of Moll: Histochemical and Ultrastructural Characterization of the Glands of Moll in the Human Eyelid. J. Invest. Dermatol., 121, 28-36.

Szeder, V., Grim, M., Halata, Z., Sieber-Bluma, M. (2003) Neural crest origin of mammalian Merkel cells. Dev. Biol., 253, 258-263.

Temple, S. (2001) The development of neural stem cells. Nature, 414, 112–117.

Towler, D.A., Gelberman, R.H. (2006) The alchemy of tendon repair: a primer for the (S)mad scientist. J. Clin. Invest., 116, 863-866.

Valentine, J. W. (2003) Cell types, numbers, and body plan complexity. In Keywords and Concepts in Evolutionary Developmental Biology (eds. B. K. Hall and W. M. Olson), pp. 35-53. Harvard University Press, Cambridge.

Vickaryous, M.K., Hall, B.K. (2006) Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. Biol. Rev. 81, 425–455.

5.2. Supplementary information

Table II. Vertices of the NHCD and the day of the cell appearance. The asterisk is the estimate appearance day.

Vertices	Name	Day
1	Fertilized egg	0
2	Blastocyst	5
3	Inner cell mass	6.5
4	Germ cell precursor	11
5	Trophoblast or outer cell mass	6.5
6	Cytotrophoblast	8.5
7	Syncytiotrophoblast	12
8	Hypoblast	8.5
9	Extraembryonic endoderm	12
10	Yolk sac	14
11	Epiblast	8.5
12	Amniotic ectoderm	12
13	Primitive ectoderm	12
14	Primitive streak	14
15	Extraembryonic mesoderm	15
16	Embryonic mesoderm	15
17	Embryonic endoderm	15
18	Embryonic ectoderm	15
19	Presomite	16
20	Neural plate	18
21	Stomodeum	18*
22	Blood island	19
23	Urogenital ridge	20
24	Primitive gut	20
25	Neural groove	19
26	Neural fold	20
27	Somite	20
28	Cardiac progenitor cell	20
29	Rathke’s pouch	20
30	Neural tube	21
31	Segmented intermediate mesoderm (pronephric system)	21
32	Foregut	21
33	Hemangioblast	21
34	Sclerotome	22
35	Myotome	22
36	Neural crest precursor	22
37	Optic groove	22
38	Neuroepithelial cell	22
39	Adrenal cortex	22
40	Otic placode	22
41	Cranial neural crest cell	23*
42	Trunk neural crest cell	23*
43	Diencephalic precursor	23
44	Primitive neuroblast	23*
45	Lung precursor	23
46	Hematopoietic precursor	23*
47	Nephrogenic cord	24*
48	Pharyngeal gut	24

49	Optic vesicles	24
50	Lens placode	24
51	Apolar neuroblast	24*
52	Germ cell	25
53	Nephric duct	25
54	Midgut	25
55	Hindgut	25
56	Primitive intestine	25
57	Liver precursor	25
58	Esophagus precursor	25
59	Primitive stomach	25
60	Pancreas precursor	25
61	Pharyngeal arch	25
62	Pharyngeal pouch	25
63	Schwann cell precursor	25
64	Optic cup	25
65	Bipolar neuroblast	25*
66	Cloaca	26
67	Pharyngeal precursor	26
68	Laryngeal precursor	26
69	Hemocytoblast	26*
70	Multipolar neuroblast	26*
71	Nasal placode	26
72	Maxillary prominences	26
73	enteric nervous system precursor	26*
74	Primitive tongue	27
75	Dermatome	27
76	Retinal pigment epithelium	27*
77	Cerebellar cortex precursor	27
78	Amygdaloid nuclei precursor	27
79	Mesonephric (wolffian) duct	28
80	Metanephric blastema	28
81	Retinal progenitor cell	28*
82	Lens vesicle	28
83	Amygdaloid fiber	28
84	Mesenchymal progenitor	28
85	Tracheal precursor	29
86	Bronchial precursor	29
87	Mono-myeloblast	29*
88	Enteric neuroblast	29
89	Sympathetic neuroblast	29
90	Sensory neuroblast	29
91	Parasympathetic neuroblast	29
92	Gonad precursor	30
93	Parathyroid gland precursor	30
94	Thymus precursor	30
95	Cone photoreceptor precursor	30
96	Rod photoreceptor precursor	30
97	Otocysts	30
98	Ureteric bud	31
99	Epithelial progenitor	31
100	Angioblast progenitor	31
101	Smooth muscle progenitor	31
102	Stromal cell progenitor	31
103	Paramesonephric (mullerian) ducts	31
104	Descending colon precursor	32
105	Small intestine	32

106	Ascending colon precursor	32
107	Promonocyte	32*
108	Neuron, enteric, interneuron	32
109	Neuron, enteric, motor to gastroenteropancreatic endocrine cell	32
110	Neuron, enteric, motor, excitatory	32
111	Neuron, enteric, motor, inhibitory	32
112	Neuron, enteric, secretomotor/vasomotor	32
113	Neuron, enteric, sensory	32
114	Neuron, enteric, stubby	32
115	Neuron, sympathetic, adrenergic	32
116	Neuron, sympathetic, B-fibre	32
117	Neuron, sympathetic, C-fibre	32
118	Neuron, sympathetic, cholinergic	32
119	Neuron, sensory, A-beta IB	32
120	Neuron, sensory, A-beta IIB	32
121	Neuron, sensory, A-delta cold fibre	32
122	Neuron, sensory, A-delta type I	32
123	Neuron, sensory, A-delta type II	32
124	Neuron, sensory, C-fibre non-peptidergic	32
125	Neuron, sensory, C-fibre peptidergic	32
126	Neuron, sensory, mechanoreceptor	32
127	Neuron, sensory, proprioceptor	32
128	Neuron, sensory, warm fibre	32
129	Neuron, parasympathetic, cholinergic	32
130	Neuron, parasympathetic, nonadrenergic/non-cholinergic	32
131	Cone photoreceptors, blue sensitive	32
132	Cone photoreceptors, green sensitive	32
133	Cone photoreceptors, red sensitive	32
134	Rod photoreceptor cell	32
135	Pituitary anterior lobe precursor	32
136	Pituitary intermediate lobe precursor	32
137	Ependymoblast	32*
138	Nasal pit	32
139	Lacrimal gland precursor	32
140	Urogenital sinus	33
141	Pars tuberalis	33
142	Urinary bladder precursor	34
143	Primitive oral cavity	35
144	Monocyte	35*
145	Hypothalamic sulcus	35
146	Gliablast	35
147	Medulla oblongata precursor	35
148	Urethra precursor	36
149	Primitive oral cavity	36*
150	Chondroblast	36
151	Ureter precusor	36
152	Primitive renal pelvis	36
153	Infundibulum	36
154	Cell of stratified squamous epithelium of oral cavity	37*
155	Semicircular canal	37
156	Oogonia	38
157	Microglia	38
158	Optic stalk	38
159	Pineal (epiphysis) precursor	38
160	Pituitary posterior lobe precursor	38
161	Nasal prominence	38
162	Dental lamina	38

163	Choroid	38
164	Transitory chondrocyte	38*
165	Urethra basal cell	40*
166	Chondrocyte, hyaline cartilage	40
167	Mammillary body	40
168	Thalamus precursor	40
169	Hypothalamus precursor	40
170	Cerebral cortex precursor	40
171	Spinal cord precursor	40
172	Substantia nigra precursor	40
173	Olfactory bulb precursor	40
174	Cochlear duct precursor	40
175	Major calyx precursor	40
176	Posterior semicircular canal	40
177	Lateral semicircular canal	40
178	Superior semicircular canal	40
179	Hippocampus precursor	42
180	Olfactory neurons precursor	42
181	Mesothelial progenitor	42
182	Sclera	42
183	Primary lens fibre	42
184	Anal canal precursor	43
185	Testis precursor	45
186	Ovary precursor	45
187	Optic nerve	45
188	Periderm	45
189	Choroid plexus	45
190	Neuron, Purkinje cell	45
191	Endolymphatic sac precursor	45
192	Pupil precursor	45
193	Triangular primary palate	45
194	Tubotympanic recess	45
195	Neuron, olfactory	46
196	Neuron, tufted cell	46
197	Secondary lens fibre	47
198	Neuron, Retzius-Cajal cell	47
199	Basal layer	50*
200	Leydig cell of testis	52
201	tubules of the rete testis	52
202	dental bud	52
203	Zona fasciculata cell	52
204	Zona glomerulosa cell	52
205	Zona reticularis cell	52
206	Neuron-specific enolase	52
207	Uterine tube	55
208	Uterine canal	55
209	Bronchiolar precursor	40
210	Anterior lens epithelium	42
211	Collecting tubules	45
212	Thyroid gland progenitor	27
213	Nephroi of mesonephro	28*
214	Nephroi of metanephro	28*
215	Oxyntic pit-gland	30*
216	Mucous pit-gland unit of pyloric antrum	30*
217	Thyroid follicular cell	30*
218	Pharyngeal epithelial cell	30*
219	Basal cell of tongue	30*

220	Cell of stratified squamous epithelium of tongue	34*
221	Ultimobranchial body	30*
222	Parafollicular cell	34*
223	Thymic epithelial cell	34*
224	Endothelial precursor	26*
225	Endothelial cell, continuous	31*
226	Endothelial cell, CNS	31*
227	Endothelial cell, fenestrated	31*
228	Endothelial cell, lymphatic	31*
229	Endothelial cell, corneal	31*
230	Endothelial cell, splenic	31*
231	Lymphoid precursor	29*
232	Erythroid/megakacariocyte precursor	29*
233	Single heart tube	26*
234	Muscle cell, cardiac fibre	32*
235	Muscle cell, cardiac ordinary	32*
236	Muscle cell, cardiac nodal cell	32*
237	Enteric (astrocyte-like) glial cell	29*
238	Immature schwann cells	27*
239	Pro-myelin schwann cell	30*
240	Myelinating schwann cell	34*
241	Non-myelinating schwann cell	34*
242	Duct cell, gall bladder	40*
243	cell of stratified squamous epithelium of urethra	45*
244	Pre-pit cell precursor	35*
245	Pre-pit cell	38*
246	Differentiating pit cell	41*
247	Pit cell	45*
248	Pre-parietal cell precursor	35*
249	Stage-1 parietal cell	38*
250	Stage-2 parietal cell	40*
251	Stage-3 parietal cell	42*
252	Parietal cell	45*
253	Pre-neck cell precursor	35*
254	Pre-neck cell	38*
255	Neck cell	40*
256	Pre-zymogenic cell	42*
257	Zymogenic cell	45*
258	Mottled-granule cell	33*
259	Mixed-granule cell	36*
260	Pre-gland cell	39*
261	Differentiating gland cell	42*
262	Gland cell	45*
263	Tracheal basal cell	35*
264	Tracheal parabasal cell	40*
265	Bronchial mucous-gland duct cells	35*
266	Bronchial basal cell	35*
267	Traqueal mucous-gland duct cells	35*
268	Traqueal secretory cell	35*
269	Bronchial secretory cell	35*
270	Ciliated cell	40*
271	Epithelial cell, respiratory, brush cell	40*
272	Bronchial parabasal cell	40*
273	Pre-absorptive cell	40*
274	Differentiating absorptive cell	45*
275	Pre-goblet	40*
276	Differentiating goblet cell	45*

277	Pre-Paneth cell	40*
278	Differentiating Paneth cell	45*
279	Pre-M cell	40*
280	M cell	45*
281	Pre-cavuolated cell	40*
282	Cavuolated cell	45*
283	Pre-vacuolated cell	36*
284	Vacuolated cell	40*
285	Columnar cell	45*
286	Pre-deep crypt secretory cell	36*
287	Differentiating deep crypt secretory cell	40*
288	Deep crypt secretory cell	45
289	Duct cell precursor	35*
290	Duct cell	45*
291	Acinar cell precursor	35*
292	Acinar cell	45*
293	Endocrine lineage precursor	35*
294	Alpha cell	45*
295	Beta cell	45*
296	Delta cell	45*
297	PP cell	45*
298	Hepatoblast	35*
299	Oval cell	45*
300	Hepatocyte	45*
301	Cholangiocyte	45*
302	Principal (chief) cell	40*
303	Oxyphil cell	40*
304	Cortical epithelial progenitor cell	40*
305	Cortical epithelial cell type I	45*
306	Cortical epithelial cell type II	45*
307	Medullary epithelial progenitor cell	40*
308	Medullary epithelial cell type I	45*
309	Medullary epithelial cell type II	45*
310	Plasmoblast	35*
311	Plasmocyte	45*
312	NK precursor	35*
313	NK cell	45*
314	Lymphoid dendritic cell precursor	35*
315	Lymphoid dendritic cell	40*
316	Thymocyte DN premature	35*
317	Thymocyte DN delayed	40*
318	Thymocyte DP premature	45*
319	Erythroblast precursor	35*
320	Proerythroblast	37*
321	Basophilic erythroblast"	39*
322	Erythroblast	41*
323	Late eritroblast	43*
324	Normoblast	45*
325	Megakaryoblast	35*
326	Megakaryocyte	40*
327	Platelet	45*
328	Promyelocyte	35*
329	Early neutrophilic myelocyte	38*
330	Neutrophilic myelocyte	41*
331	Neutrophilic metamyelocyte	45*
332	Early basophilic promyelocyte	38*
333	Basophilic myelocyte	41*

334	Basophilic metamyelocyte	45*
335	Early eosinophilic promyelocyte	38*
336	Eosinophilic myelocyte	41*
337	Eosinophilic metamyelocyte	45*
338	Immature mast cell	40*
339	Mature mast cell	45*
340	Immature macrophage	38*
341	Kupffer cell	38*
342	Langerhans cell	38*
343	chondroclast	38*
344	Multinucleated osteoclast	38*
345	Mononucleated osteoclast	38*
346	Macrophage-antigen-positive microglia	38*
347	Macrophage-antigen-negative microglia	38*
348	Plasmacytoid dendritic cell	38*
349	Chondrocyte, elastic cartilage	40*
350	Chondrocyte, fibrocartilage	40*
351	Adipoblast	40*
352	Preadipocyte	45*
353	Osteoprogenitor	40*
354	Transitory osteoblast	45*
355	Fibroblast progenitor	40*
356	Transitory fibroblast	45*
357	Tenocyte progenitor	40*
358	Tenocyte	45*
359	Transitory stromal cell	40*
360	Unique micro-niche	45*
361	Pericyte	40*
362	Nucleus pulposus cell	40*
363	Annulus fibrous cell	45*
364	Myoblast precursor	30*
365	Myoblast	35*
366	Satellite cell	35*
367	Side population cell	35*
368	Muscle cell, skeletal, intermediate	40*
369	Muscle cell, skeletal, red cell (slow)	40*
370	Muscle cell, skeletal, white cell (fast)	40*
371	Myotube	40*
372	Spindle precursor	45*
373	Muscle spindle, nuclear bag	45*
374	Muscle spindle, nuclear chain	35*
375	Angioblast	40*
376	Cell in the wall of the dorsal aorta	40*
377	Corticotrophe	40*
378	Gonadotroph	40*
379	Thyrotrope	40*
380	Somatotroph/lactotroph precursor	40*
381	Somatotrophe	45*
382	Lactotrophe	45*
383	Melanotrope cell	40*
384	Oxytocin-secreting cell	45*
385	Vasopressin-secreting cell	45*
386	Lacrimal gland cell	45*
387	Esohagus basal cell	35*
388	Prickle cell	45*
389	Squamous cell	45*
390	Premature pro-B cell	35*

391	Delayed pro-B cell	38*
392	Big pre-B cell	41*
393	Small pre-b cell	44*
394	Immature b cell	45*
395	Optic nerve fiber	60
396	Astrocyte, protoplasmic (velate)	60
397	Astrocyte, fibrous	60
398	Lip	60
399	Secondary palate	60
400	Scala tympani	63
401	Outer dental epithelium	65
402	Inner dental epithelium	65
403	Dental papilla	65
404	Archicortex precursor	65
405	Vestibular apparatus precursor	65
406	Perilymphatic precursor	65
407	Primary oocyte	75
408	odontoblast	75
409	Corium	60*
410	Dermal papillae	75
411	Stellate reticulum	75
412	Ameloblast	75
413	Gastroenteropancreatic glucagon secreting (A) cell	60*
414	Vaginal basal cell	80
415	Ganglion precursor	80
416	Horizontal precursor	80
417	Amacrine precursor	80
418	Bipolar precursor	80
419	Intermediate layer	80
420	Neuron, nigral dopaminergic cell	80
421	Littré gland precursor	95
422	Principal cell of collecting duct	95
423	intercalated cell of collecting duct	95
424	Podocyte	95
425	Parietal	95
426	Proximal tubule	95
427	Distal tubule	95
428	Convolute tubule	95
429	Macula densa	95
430	Thick loop of Henle	95
431	Thin limb loop of Henle	95
432	Glomerular endothelium	95
433	Peritubular endothelium	95
434	Mesangial	95
435	Vascular	95
436	Ureteric	95
437	Detrusor	95
438	Cortical interstitial cell	95
439	Medullary interstitial cell	95
440	Epigenital tubule	95
441	Paragenital tubule	95
442	Sertoli cell	100
443	Follicular cell	100
444	Ductus deferen	100
445	Neuron, granule cell	100
446	Hair bud	105
447	Oligodendrocyte precursor	108

448	Cochlear nuclei precursor	108
449	Efferent ductules	110
450	Epididymal precursor	110
451	Seminal vesicle precursor	110
452	Spinous layer	110
453	Granular layer	110
454	Horny layer	110
455	Paraurethral gland precursor	110
456	Paradidymis	110
457	Appendix epididymis	110
458	Oligodendrocyte, interfascicular	115
459	Oligodendrocyte, satellite	115
460	Dental pulp precursor	115*
461	Pinealoblast	128*
462	Hair matrix cell	135*
463	Vaginal fornice	135
464	Sinovaginal bulb	60
465	Uterus precursor	60
466	Mammary precursor	75
467	Cuboidal follicular cell	100*
468	Primary follicle	120*
469	Secondary follicle	135*
470	Theca folliculi	100*
471	Theca interna cell	120*
472	Theca externa cell	120*
473	Myofibroblast	50*
474	Juxtaglomerular cell	60*
475	Anal canal basal cell	60*
476	Cell of stratified squamous epithelium of anal canal	75*
477	Prostate gland precursor	80
478	Basal cell	100*
479	Specialized basal cell	100*
480	Luminal cell	120*
481	Specialized luminal cell	130*
482	Neuroendocrine cell	100*
483	Cell of stratified squamous epithelium of vagina	100*
484	Epithelial cell of endometrium of uterus	100*
485	Ciliated cell of endometrium of uterus	100*
486	Seminiferous tubules	100*
487	Epididymal principal cell	125*
488	Epididymal basal cell	125*
489	Secretory seminal vesicle cell	125*
490	Duct cell of seminal vesicle	125*
491	Nonciliated cell of efferent ductules	125*
492	Ciliated cell of Efferent ductules	125*
493	Basal cell of urinary epithelium	100*
494	Cell of urinary epithelium	125*
495	Clara cell	60*
496	Pulmonary neuroendocrine precursor	60*
497	Absorptive cell	50*
498	Goblet cell	50*
499	Paneth cell	50*
500	Mature b cell	50*
501	Lymphocyte IgA B	55*
502	Lymphocyte IgB B	55*
503	Lymphocyte IgD B	55*
504	Lymphocyte IgE B	55*

505	Lymphocyte IgG B	55*
506	Lymphocyte IgM B	55*
507	Thymocyte DP delayed	50*
508	Virgin CD4 T cell	55*
509	Memory CD4 T cell	75*
510	Th1 cell	75*
511	Th2 cell	75*
512	Th3 cell	75*
513	CD4 NK1.1 T cell	75*
514	CD8 T cell	60*
515	Memory CD8 T cell	75*
516	Effective CD8 T cell	90*
517	Late normoblast	50*
518	Erythrocyte	55*
519	Neutrophilc leukocyte	60*
520	Basophilc leukocyte	60*
521	Eosinophil leukocyte	60*
522	Immature myeloid dendritic cell	50*
523	Mature myeloid dendritic cell	60*
524	Macrophage, alveolar	75*
525	Macrophage, connective tissue	75*
526	Macrophage, multinucleated	75*
527	Macrophage, serous cavity	75*
528	Macrophage, splenic	75*
529	Osteoblast	55*
530	Osteocyte	65*
531	Fibroblast	55*
532	Stromal cell	55*
533	Adipose cell, brown	60*
534	Adipose cell, white	60*
535	Lipocyte of liver	60*
536	Melanoblast	80*
537	Melanocyte	120*
538	Satellite cell of ganglia	100*
539	Basal cell of cornea	75*
540	Cell of stratified squamoous ephitelium of cornea	100*
541	Myoepithelial cell, iris	75*
542	Squamous cell, pia-arachnoid	100*
543	Small intensely fluorescent precursor	100*
544	Small intensely fluorescent (SIF) cell, type I	125*
545	Small intensely fluorescent (SIF) cell, type II	125*
546	Adrenal medulla precursor	75*
547	Chromaffin cell, epinephrine secreting	100*
548	Chromaffin cell, norepinephrine secreting	100*
549	Epithelial cell precursor (eye)	60*
550	Ciliary pigmented epithelial cell (eye)	100*
551	Ciliary unpigmented epithelial cell (eye)	100*
552	Sebaceous gland cell precursor	120*
553	Sebaceous gland cell	135*
554	Fallopian tube precursor	75*
555	Secretory fallopian tube cell	100*
556	Sweat gland precursor	120*
557	Eccrine sweat gland precursor	130*
558	Gland cell, sweat, eccrine dark	135*
559	Gland cell, sweat, eccrine clear	135*
560	Apocrine sweat gland precursor	130*
561	Gland cell, sweat, apocrine	135*

562	Duct cell, nonstriated	135*
563	Sphincter pupillae muscle	80*
564	Dilator pupillae muscle	80*
565	Smooth muscle of uterus precursor	130*
566	Smooth muscle of uterus	135*
567	Smooth muscle of gut precursor	60*
568	Smooth muscle of gut	80*
569	Smooth muscle of the mammary gland precursor	130*
570	Smooth muscle of the mammary gland	135*
571	Keratinocyte precursor	120*
572	Keratinocyte	135*
573	Hyalocyte, vitreous body of eye	100*
574	Ependymal cell	50*
575	Tanycyte	100*
576	Neuron, thalamic relay	100*
577	Neuron, thalamic reticular	100*
578	Bergmann cell	100*
579	Olfactory cortex precursor	50*
580	Zona fasciculata and reticularis cell, mineralocorticoid-secreting cell	60*
581	Zona fasciculata and reticularis cell, glucocorticoid-secreting cell	60*
582	Serosal progenitor	60*
583	Serosal cell	80*
584	Salivary gland cell precursor	75*
585	Mucous cell	100*
586	Serous cell	100*
587	Intercalated duct cell of salivary gland	100*
588	Striated duct cell of salivary gland	100*
589	Neostratum precursor	60*
590	Basal cell of olfactory epithelium	75*
591	Epithelial cell, olfactory, olfactory cell	100*
592	Epithelial cell, olfactory, supporting	100*
593	Epithelial cell, olfactory, sustentacular cell	100*
594	Pia cell	50*
595	Stria vascularis basal cell	75*
596	Stria vascularis intermediate cell	75*
597	Stria vascularis marginal cell	75*
598	Supporting cell precursor, organ of corti	100*
599	Supporting cell, organ of corti	135*
600	Phalangeal precursor	125*
601	Phalangeal cell, inner supporting	135*
602	Phalangeal cell, outer supporting	135*
603	Pillar cell precursor	125*
604	Pillar cell, inner supporting	135*
605	Pillar cell, outer supporting	135*
606	Supporting border cell	135*
607	Hensen cell	135*
608	Deiter cell	135*
609	Sensory hair precursor (organ of Corti)	100*
610	Hair cell, organ of Corti, inner	135*
611	Hair cell, organ of Corti, outer	135*
612	Sensory acceleration and gravity cell precursor	100*
613	Hair cell, type I	135*
614	Hair cell, type II	135*
615	Supporting cell	100*
616	Light cell	100*
617	Vestibular apparatus membrane cell	100*
618	Dark cell precursor	135*

619	Dark cell, crista ampullaris	135*
620	Dark cell, endolymphatic sac	135*
621	Columnar precursor	100*
622	Columnar cell with microvilli	135*
623	Columnar cell without microvilli	135*
624	Squamous cell, endolymphatic sac lining	100*
625	Perilymphatic precursor	65
626	Squamous cell, perilymphatic space lining	135*
627	Stellate cell, perilymphatic space lining	135*
628	Crus ampullare	135*
629	Crus nonampullare	135*
630	Pituicyte	60*
631	Neuron motor precursor	60*
632	Iridopupillary membrane	100*
633	Substantia propria of the cornea	100*
634	Pars ceca retinae	100*
635	Cuboidal mesothelial cell	100*
636	Squamous-like mesothelial cell	100*
637	Ureter epithelial cell	100*
638	Minor calyx	120*
639	Neuron, ganglion G3	140
640	Neuron, ganglion G4	140
641	Neuron, ganglion G5	140
642	Neuron, ganglion G7	140
643	Neuron, ganglion G8	140
644	Neuron, ganglion G10	140
645	Neuron, ganglion G11	140
646	Neuron, ganglion G12	140
647	Neuron, ganglion G16	140
648	Neuron, ganglion G17	140
649	Neuron, ganglion G19	140
650	Neuron, ganglion G20	140
651	Neuron, ganglion G21	140
652	Neuron, ganglion G22	140
653	Neuron, ganglion G23	140
654	Neuron, ganglion M	140
655	Neuron, ganglion P1	140
656	Neuron, ganglion P2	140
657	Neuron, horizontal type HI	140
658	Neuron, horizontal type HII	140
659	Neuron, horizontal type HIII	140
660	Neuron, amacrine cholinergic stratum 4	140
661	Neuron, amacrine semilunar type 1	140
662	Neuron, amacrine semilunar type 2	140
663	Neuron, amacrine small-field diffuse	140
664	Neuron, amacrine spiny	140
665	Neuron, amacrine stellate-varicose	140
666	Neuron, amacrine thorny type 1	140
667	Neuron, amacrine thorny type 2	140
668	Neuron, amacrine tristratified	140
669	Neuron, amacrine type A1	140
670	Neuron, amacrine type A2	140
671	Neuron, amacrine type A3	140
672	Neuron, amacrine type A4	140
673	Neuron, amacrine type A5	140
674	Neuron, amacrine type A8	140
675	Neuron, amacrine type A12	140

676	Neuron, amacrine type A13	140
677	Neuron, amacrine type A14	140
678	Neuron, amacrine type A17	140
679	Neuron, amacrine type A18	140
680	Neuron, amacrine type AII	140
681	Neuron, amacrine wavy	140
682	Neuron, amacrine wiry	140
683	Neuron, amacrine wooly diffuse	140
684	Neuron, bistratified giant bipolar	140
685	Neuron, blue cone bipolar cell type a	140
686	Neuron, blue cone bipolar cell type b	140
687	Neuron, diffuse cone bipolar cell type a	140
688	Neuron, diffuse cone bipolar cell type b	140
689	Neuron, rod bipolar cell	140
690	Neuron, flat midget bipolar	140
691	Neuron, giant diffuse bipolar	140
692	Neuron, invaginating midget bipolar	140
693	Hair shaft cell	165
694	Epithelial root sheath	165
695	Dental pulp cell	165
696	Follicular papilla	165
697	Neuron, callosal	170
698	Neuron, double bouquet cell (DBC)	185
699	Neuron, fusiform cell	190
700	Neuron, immature basket cell	205
701	Neuron, basket cell type I	210
702	Neuron, basket cell type II	210
703	Neuron, basket cell type III	210
704	Pinealocyte	220
705	Epithelial precursor cell	225
706	Ductal precursor	225
707	Inner root sheath	240
708	Outer root sheat"	240
709	Dermal sheat	240
710	Alveolar precursor	140
711	Pneumocyte I	220*
712	Pneumocyte II	200*
713	Muller cell	140*
714	Neuron, interplexiform	140*
715	Hair shaft, medulla	220*
716	Hair shaft, cortex	220*
717	Hair shaft, hair cuticle	220*
718	Hair cell, external	200*
719	Neuron, axo-axonic	240*
720	Neuron, back-projection	240*
721	Neuron, Betz cell	240*
722	Neuron, bilaminar	240*
723	Neuron, bistratified	240*
724	Neuron, cell of Martinotti	240*
725	Neuron, chandelier	240*
726	Neuron, hippocampal-septal	240*
727	Neuron, interneurone specific cell precursor	220*
728	Neuron, interneurone specific cell I (IS-I)	240*
729	Neuron, interneurone specific cell II (IS-II)	240*
730	Neuron, interneurone specific cell III (IS-III)	240*
731	Neuron, lacunosaum-moleculare perforant path associated	240*
732	Neuron, lacunosaum-moleculare-radiatum perforant path associated	240*

733	Neuron, neurogliaform cell	240*
734	Neuron, Schaffer collateral associated	240*
735	Neuron, stratum oriens, lacunosummolaculare (O-LM)	240*
736	Neuron, stellate cell precursor	240*
737	Neuron, type I stellate cell	240*
738	Neuron, type II stellate cell	240*
739	Neuron, Golgi cell	240*
740	Neuron, unipolar brush cell (UBC)	240*
741	Neuron, Renshaw cell	240*
742	Neuron, spinal Ia interneuron	240*
743	Neuron, spinal motor	240*
744	Neuron, CA1 oriens alveus interneuron	240*
745	Neuron, CA1 pyramidal	240*
746	Neuron, CA3 pyramidal	240*
747	Neuron, mitral cell	240*
748	Neuron, periglomerular cell	240*
749	Neuron, olfactory bulb granule cell	240*
750	Neuron, olfactory cortex interneuron (deep)	240*
751	Neuron, olfactory cortex interneuron (superficial)	240*
752	Neuron, olfactory cortex pyramidal cell	240*
753	Myoepithelial cell	225*
754	Luminal cell	225*
755	Alveolar cell	200*
756	Neuron, striatal leptodendritic	240*
757	Neuron, striatal microneuron	240*
758	Neuron, striatal spidery	240*
759	Neuron, striatal spiny	240*
760	Neuron, amacrine cholinergic stratum 2	140*
761	Neuron, pyramidal cell	240*
762	Neuron, neostriatal cholinergic interneuron	240*
763	Neuron, octopus cell	240*
764	Neuron, antenniform cell	240*
765	Neuron, clavate cell	240*
766	Neuron, elongate cell	240*
767	Neuron, globular bushy cell	240*
768	Neuron, giant cell	240*
769	Neuron, large spherical bushy cell	240*
770	Neuron, small spherical bushy cell	240*
771	Neuron, dentate granule cell	240*
772	Neuron, motor, A-alpha (IA)	240*
773	Neuron, motor, A-gamma	240*
774	Pars iridica retinae	160*
775	Inner layer of the iris precursor	200*
776	Outer pigmented layer of iris	240*
777	Inner unpigmented layer of iris	240*
778	Pars ciliaris retinae	200*
779	Ciliary body precursor	240*
780	Hymen	200*
781	Littré gland cell	140*
782	Paraurethral gland cell	140*
783	Secondary oocyte	
784	Polar body	
785	Mature oocyte	
786	Spermatogonial stem cells	
787	Type A spermatogonia	
788	Type B spermatogonia	
789	Primary spermatocyte	

790 Secondary spermatocyte
791 Spermatid
792 Spermatozoa
793 Argentaffin cell
794 Gastroenteropancreatic enterochromaffin- like (ECL) cell
795 Gastroenteropancreatic endocrine (P) cell
796 Gastroenteropancreatic gastrin secreting (G) cell
797 Gastroenteropancreatic X cell
798 Gastroenteropancreatic CCK secreting cell
799 Gastroenteropancreatic glucagon-dependent insulinotropic peptide secreting (GIP) cell
800 Gastroenteropancreatic glucagon-like peptide-2 secreting (GLP-2) cell
801 Gastroenteropancreatic neurotensin secreting (N) cell
802 Gastroenteropancreatic peptide tyrosine secreting (PYY) cell
803 Gastroenteropancreatic secretin secreting (S) cell
804 Gastroenteropancreatic very large (VL) cell
805 Calcitonin-producing cell
806 Carotid body type I cell
807 Parafollicular (C) cell
808 Synovial A cell
809 Synovial B cell
810 Smooth muscle of sweat gland
811 Carotid body type I cell
812 Synovial precursor
813 Cuticle cell
814 Hair cell, Henle’s layer
815 Hair cell, Huxley’s layer
816 Cerumious gland cell precursor
817 Cerumious gland cell
818 Moll gland cell precursor
819 Moll gland cell
820 von Ebner gland precursor
821 von Ebner gland cell
822 Bowman's precursor cell
823 Bowman’s dark cell
824 Bowman’s light cell
825 Interstitial cell of Cajal precursor
826 Interstitial cell of Cajal, myenteric region (IC-MY)
827 Interstitial cell of Cajal, submucosal region (IC-SM)
828 Interstitial cell of Cajal, deep muscularis plexus region (IC-DMP)
829 Interstitial cell of Cajal, intramuscular region (IC-IM)
830 Merkel cell precursor
831 Merkel cell
832 Secretory cell, endorphin
833 Gastroenteropancreatic motilin secreting (M) cell
834 Pulmonary (P) endocrine cell, bombesin
835 Gastroenteropancreatic gastrin/CCK secreting cell
836 Taste bud precursor
837 Type I taste bud cell
838 Type II taste bud cell
839 Perioddontal ligament
840 Cementoblast
841 Odontoclast
842 Cementoclast
843 Bulbourethral gland precursor
844 Bulbourethral gland cell
845 Bartholin’s gland precursor
846 Bartholin’s gland cell

847	Brunner’s gland precursor	
848	Brunner’s gland cell	
849	Purkinje fibre cell	
850	Pituitary folliculo-stellate cell	
851	Planum semilunatum cell	
852	Boettcher cell	
853	Claudis cell	
854	Interdental cell	
855	Fibroblastic reticular cell	
856	Littorial cell	
857	Conjunctival epithelium precursor	
858	Epithelial cell, conjunctival	
859	Cementocyte	
860	Odontocyte	
861	Primitive tympanic cavity	
862	Auditory tube	
863	Malleus	
864	Incus	
865	Stapes	
866	Auricle precursor	
867	Antihelix	
868	Concha	
869	Helix	
870	Cymba conchae	
871	Tragus	
872	Antitragus	
873	Pituitary cell, intermediate, MSH secreting	50*

Table III. Edges of the NHCD and the day of the cell appearance. The asterisk is the estimate appearance day.

i	j	Day	Reference
1	2	5	Kirschstein and Skirboll, 2001
2	3	6.5	Kirschstein and Skirboll, 2001
2	5	6.5	Kirschstein and Skirboll, 2001
5	6	8.5	Kirschstein and Skirboll, 2001
6	7	12	Kirschstein and Skirboll, 2001
3	8	8.5	Kirschstein and Skirboll, 2001
8	9	12	Kirschstein and Skirboll, 2001
9	10	14	Kirschstein and Skirboll, 2001
3	11	8.5	Kirschstein and Skirboll, 2001
11	12	12	Kirschstein and Skirboll, 2001
11	13	12	Kirschstein and Skirboll, 2001
13	14	14	Kirschstein and Skirboll, 2001
13	18	15	Kirschstein and Skirboll, 2001
14	15	15	Kirschstein and Skirboll, 2001
14	16	15	Kirschstein and Skirboll, 2001
14	17	15	Kirschstein and Skirboll, 2001
11	4	11	Kirschstein and Skirboll, 2001; Sell, 2004
16	19	16	Sadler, 2004
18	20	18	Paxinos and Mai, 2004; Sadler, 2004
18	21	18*	Sadler, 2004
16	22	19	Sadler, 2004
16	23	20	Sadler, 2004
17	24	20	Sadler, 2004
20	25	19	Paxinos and Mai, 2004; Sadler, 2004
25	26	20	Paxinos and Mai, 2004; Sadler, 2004
19	27	20	Sadler, 2004
16	28	20	Sadler, 2004
21	29	20	Sadler, 2004
26	30	21	Paxinos and Mai, 2004; Sadler, 2004
26	31	21	Horster et al., 1999; Sadler, 2004
24	32	21	Sadler, 2004
22	33	21	Sadler, 2004; Kirschstein and Skirboll, 2001
27	34	22	Sadler, 2004
27	35	22	Sadler, 2004
18	36	22	Paxinos and Mai, 2004; Sadler, 2004
20	37	22	Sadler, 2004; Sell, 2004
30	38	22	Paxinos and Mai, 2004; Sadler, 2004
23	39	22	Sadler, 2004
18	40	22	Sadler, 2004
36	41	23*	Santagati and Rijli, 2003
36	42	23*	Santagati and Rijli, 2003
38	43	23	Sadler, 2004
38	44	23*	Sadler, 2004
32	45	23	Sadler, 2004
33	46	23*	Kirschstein and Skirboll, 2001; Sadler, 2004
23	47	24*	Horster et al., 1999
24	48	24	Sadler, 2004
37	49	24	Sell, 2004
49	50	24	Sadler, 2004; Sell, 2004
44	51	24*	Sadler, 2004
4	52	25	Kirschstein and Skirboll, 2001; Sadler, 2004; Sell, 2004

23	53	25	Horster et al., 1999; Sadler, 2004
24	54	25	Sadler, 2004
24	55	25	Sadler, 2004
32	56	26	Sadler, 2004
54	56	26	Sadler, 2004
32	57	25	Sadler, 2004
32	58	25	Sadler, 2004
32	59	25	Sadler, 2004
32	60	25	Sadler, 2004
48	61	25	Sadler, 2004
48	62	25	Sadler, 2004
41	63	25	Sadler, 2004; Santagati and Rijli, 2003
42	63	25	Sadler, 2004; Santagati and Rijli, 2003
49	64	25	Sadler, 2004; Sell, 2004
51	65	25*	Paxinos and Mai, 2004; Sadler, 2004
55	66	26	Sadler, 2004
61	67	26	Sadler, 2004
61	68	26	Sadler, 2004
46	69	26	Alberts et al., 2002
65	70	26	Sadler, 2004
18	71	26	Sadler, 2004
61	72	26	Sadler, 2004
41	73	26*	Santagati and Rijli, 2003
61	74	27	Sadler, 2004
27	75	27	Sadler, 2004
64	76	27*	Sell, 2004
70	77	27	Paxinos and Mai, 2004; Sadler, 2004
43	78	27	Paxinos and Mai, 2004
53	79	28	Horster et al., 1999; Sadler, 2004
47	80	28	Anglani et al., 2004; Horster et al., 1999; Sadler, 2004
53	80	28	Anglani et al., 2004; Horster et al., 1999; Sadler, 2004
64	81	28*	Sell, 2004
50	82	28	Sadler, 2004; Sell, 2004
78	83	28	Paxinos and Mai, 2004
34	84	28	Sadler, 2004; Sell, 2004
45	85	29	Otto, 2002; Sadler, 2004
45	86	29	Otto, 2002; Sadler, 2004
69	87	29*	Janeway et al., 2001; Sell, 2004
73	88	29	Sadler, 2004; Vickaryous and Hall, 2006
42	89	29	Sadler, 2004
42	90	29	Sadler, 2004
41	91	29	Sadler, 2004
23	92	30	Sadler, 2004
62	93	30	Sadler, 2004
62	94	30	Sadler, 2004
81	95	30	Sadler, 2004; Sell, 2004
81	96	30	Sadler, 2004; Sell, 2004
40	97	30	Sadler, 2004
79	98	31	Anglani et al., 2004; Sadler, 2004
80	99	31	Anglani et al., 2004
80	100	31	Anglani et al., 2004
80	101	31	Anglani et al., 2004
80	102	31	Anglani et al., 2004
23	103	31	Sadler, 2004
56	104	32	Sadler, 2004
56	105	32	Sadler, 2004
56	106	32	Sadler, 2004

87	107	32*	Sell, 2004
88	108	32	Sadler, 2004; Vickaryous and Hall, 2006
88	109	32	Sadler, 2004; Vickaryous and Hall, 2006
88	110	32	Sadler, 2004; Vickaryous and Hall, 2006
88	111	32	Sadler, 2004; Vickaryous and Hall, 2006
88	112	32	Sadler, 2004; Vickaryous and Hall, 2006
88	113	32	Sadler, 2004; Vickaryous and Hall, 2006
88	114	32	Sadler, 2004; Vickaryous and Hall, 2006
89	115	32	Sadler, 2004; Vickaryous and Hall, 2006
89	116	32	Sadler, 2004; Vickaryous and Hall, 2006
89	117	32	Sadler, 2004; Vickaryous and Hall, 2006
89	118	32	Sadler, 2004; Vickaryous and Hall, 2006
90	119	32	Sadler, 2004; Vickaryous and Hall, 2006
90	120	32	Sadler, 2004; Vickaryous and Hall, 2006
90	121	32	Sadler, 2004; Vickaryous and Hall, 2006
90	122	32	Sadler, 2004; Vickaryous and Hall, 2006
90	123	32	Sadler, 2004; Vickaryous and Hall, 2006
90	124	32	Sadler, 2004; Vickaryous and Hall, 2006
90	125	32	Sadler, 2004; Vickaryous and Hall, 2006
90	126	32	Sadler, 2004; Vickaryous and Hall, 2006
90	127	32	Sadler, 2004; Vickaryous and Hall, 2006
90	128	32	Sadler, 2004; Vickaryous and Hall, 2006
91	129	32	Sadler, 2004; Vickaryous and Hall, 2006
91	130	32	Sadler, 2004; Vickaryous and Hall, 2006
95	131	32	Sadler, 2004; Vickaryous and Hall, 2006
95	132	32	Sadler, 2004; Vickaryous and Hall, 2006
95	133	32	Sadler, 2004; Vickaryous and Hall, 2006
96	134	32	Sadler, 2004; Vickaryous and Hall, 2006
29	135	32	Paxinos and Mai, 2004; Savage et al., 2003
29	136	32	Paxinos and Mai, 2004; Sadler et al., 2004
38	137	32*	Paxinos and Mai, 2004
71	138	32	Sadler, 2004
64	139	32	Sadler, 2004
66	140	33	Sadler, 2004
29	141	33	Paxinos and Mai, 2004; Savage et al., 2003
140	142	34	Sadler, 2004
21	143	35	Sadler, 2004
107	144	35*	Sell, 2004
43	145	35	Paxinos and Mai, 2004; Sadler et al., 2004
38	146	35	Kirschstein and Skirboll, 2001; Paxinos and Mai, 2004, Temple, 2001
70	147	35	Sadler, 2004; Paxinos and Mai, 2004
140	148	36	Sadler, 2004
143	149	36	Freitas, 1999
84	150	36	Sadler, 2004; Sell, 2004
98	151	36	Sadler, 2004
98	152	36	Sadler, 2004
38	153	36	Sadler, 2004
149	154	37*	Freitas, 1999
97	155	37	Sadler, 2004
52	156	38	Sadler, 2004
144	157	38	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
38	158	38	Sadler, 2004; Paxinos and Mai, 2004
43	159	38	Sadler, 2004; Paxinos and Mai, 2004
153	160	38	Sadler, 2004; Paxinos and Mai, 2004
138	161	38	Sadler, 2004
143	162	38	Nakashima and Redid, 2003; Sadler, 2004

64	163	38	Sadler, 2004; Sell, 2004
150	164	38*	Sell, 2004
148	165	40*	Freitas, 1999
164	166	40	Sadler, 2004; Vickaryous and Hall, 2006
43	167	40	Sadler, 2004
145	168	40	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
145	169	40	Paxinos and Mai, 2004
70	170	40	Sadler, 2004; Paxinos and Mai, 2004
70	171	40	Sadler, 2004; Paxinos and Mai, 2004
70	172	40	Sadler, 2004; Paxinos and Mai, 2004
70	173	40	Kirschstein and Skirboll, 2001; Paxinos and Mai, 2004
97	174	40	Sadler, 2004
152	175	40	Sadler, 2004
155	176	40	Sadler, 2004
155	177	40	Sadler, 2004
155	178	40	Sadler, 2004
70	179	42	Kirschstein and Skirboll, 2001; Sadler, 2004; Paxinos and Mai, 2004
71	180	42	Freitas, 1999; Sadler, 2004
16	181	42	Herrick and Mutsaers, 2004
64	182	42	Sadler, 2004
82	183	42	Sadler, 2004
66	184	43	Sadler, 2004
92	185	45	Sadler, 2004
92	186	45	Sadler, 2004
158	187	45	Sadler, 2004
18	188	45	Sadler, 2004
137	189	45	Sadler, 2004; Vickaryous and Hall, 2006
77	190	45	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
97	191	45	Paxinos and Mai, 2004
64	192	45	Sadler, 2004
72	193	45	Sadler, 2004
62	194	45	Sadler, 2004
180	195	46	Sadler, 2004; Vickaryous and Hall, 2006
180	196	46	Sadler, 2004; Vickaryous and Hall, 2006
183	197	47	Sadler, 2004
170	198	47	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
188	199	50*	Sadler, 2004
92	200	52	Sadler, 2004
185	201	52	Sadler, 2004
162	202	52	Nakashima and Redid, 2003; Sadler, 2004
39	203	52	Coulter, 2004; Vickaryous and Hall, 2006
39	204	52	Coulter, 2004; Vickaryous and Hall, 2006
39	205	52	Coulter, 2004; Vickaryous and Hall, 2006
171	206	52	Paxinos and Mai, 2004
103	207	55	Sadler, 2004
103	208	55	Sadler, 2004
86	209	40	Sadler, 2004; Otto, 2002
82	210	42	Sadler, 2004
175	211	45	Sadler, 2004
61	212	27	Sadler, 2004
47	213	28*	Horster et al., 1999
47	214	28*	Horster et al., 1999
59	215	30*	Sell, 2004
59	216	30*	Sell, 2004
212	217	30*	Vickaryous and Hall, 2006
67	218	30*	Sadler, 2004

74	219	30*	Freitas, 1999
219	220	34*	Freitas, 1999
62	221	34*	Sadler, 2004
221	222	34*	Sadler, 2004
94	223	34*	Blackburn and Manley, 2004
33	224	26*	Kirschstein and Skirboll, 2001
224	225	31*	Vickaryous and Hall, 2006
224	226	31*	Vickaryous and Hall, 2006
224	227	31*	Vickaryous and Hall, 2006
224	228	31*	Vickaryous and Hall, 2006
224	229	31*	Vickaryous and Hall, 2006
224	230	31*	Vickaryous and Hall, 2006
69	231	29*	Janeway et al., 2001
87	232	29*	Janeway et al., 2001
28	233	26*	Sadler, 2004; Sell, 2004
233	234	32*	Sadler, 2004
233	235	32*	Sadler, 2004
233	236	32*	Sadler, 2004
73	237	29*	Vickaryous and Hall, 2006
63	238	27*	Jessen and Mirsky, 2005
238	239	30*	Jessen and Mirsky, 2005
239	240	34*	Jessen and Mirsky, 2005
238	241	34*	Jessen and Mirsky, 2005
240	238	34*	Jessen and Mirsky, 2005
142	242	40*	Vickaryous and Hall, 2006
165	243	45*	Freitas, 1999
215	244	35*	Sell, 2004
244	245	38*	Sell, 2004
245	246	41*	Sell, 2004
246	247	45*	Sell, 2004
215	248	35*	Sell, 2004
248	249	38*	Sell, 2004
249	250	40*	Sell, 2004
250	251	42*	Sell, 2004
251	252	45*	Sell, 2004
244	249	38*	Sell, 2004
215	253	35*	Sell, 2004
253	254	38*	Sell, 2004
254	255	40*	Sell, 2004
255	256	42*	Sell, 2004
256	257	45*	Sell, 2004
253	249	45*	Sell, 2004
216	258	33*	Sell, 2004
258	259	36*	Sell, 2004
259	260	39*	Sell, 2004
260	261	42*	Sell, 2004
261	262	45*	Sell, 2004
259	245	38*	Sell, 2004
85	263	35*	Otto, 2002
263	264	40*	Otto, 2002
86	265	35*	Otto, 2002
86	266	35*	Otto, 2002
85	267	35*	Otto, 2002
85	268	35*	Otto, 2002
86	269	35*	Otto, 2002
263	270	40*	Otto, 2002
86	270	40*	Otto, 2002

265	270	40*	Otto, 2002
266	270	40*	Otto, 2002
267	270	40*	Otto, 2002
268	270	40*	Otto, 2002
269	270	40*	Otto, 2002
45	271	40*	Vickaryous and Hall, 2006
86	272	40*	Otto, 2002
105	273	40*	Sell, 2004
273	274	45*	Sell, 2004
105	275	40*	Sell, 2004
275	276	45*	Sell, 2004
105	277	40*	Sell, 2004
277	278	45*	Sell, 2004
106	275	40*	Sell, 2004
104	275	40*	Sell, 2004
104	279	40*	Sell, 2004
279	280	45*	Sell, 2004
105	279	40*	Sell, 2004
106	279	40*	Sell, 2004
104	281	40*	Sell, 2004
281	282	45*	Sell, 2004
106	283	36*	Sell, 2004
283	284	40*	Sell, 2004
284	285	45*	Sell, 2004
104	283	36*	Sell, 2004
106	286	36*	Sell, 2004
286	287	40*	Sell, 2004
287	288	45*	Sell, 2004
60	289	35*	Bardeesy and DePinho, 2002
289	290	45*	Bardeesy and DePinho, 2002
60	291	35*	Bardeesy and DePinho, 2002
291	292	45*	Bardeesy and DePinho, 2002
60	293	35*	Bardeesy and DePinho, 2002
293	294	45*	Bardeesy and DePinho, 2002
293	295	45*	Bardeesy and DePinho, 2002
293	296	45*	Bardeesy and DePinho, 2002
293	297	45*	Bardeesy and DePinho, 2002
57	298	35*	Sadler, 2004
298	299	45*	Fausto, 2004
298	300	45*	Fausto, 2004
298	301	45*	Fausto, 2004
93	302	40*	Vickaryous and Hall, 2006
93	303	40*	Freitas, 1999
223	304	40*	Blackburn and Manley, 2004
304	305	45*	Blackburn and Manley, 2004
304	306	45*	Blackburn and Manley, 2004
223	307	40*	Blackburn and Manley, 2004
307	308	45*	Blackburn and Manley, 2004
307	309	45*	Blackburn and Manley, 2004
231	310	35*	Janeway et al., 2001
310	311	45*	Janeway et al., 2001
231	312	35*	Janeway et al., 2001
312	313	45*	Janeway et al., 2001
231	314	35*	Janeway et al., 2001
314	315	45*	Janeway et al., 2001
231	316	35*	Janeway et al., 2001
316	317	40*	Janeway et al., 2001

317	318	45*	Janeway et al., 2001
232	319	35*	Janeway et al., 2001
319	320	37*	Sell, 2004
320	321	39*	Sell, 2004
321	322	41*	Sell, 2004
322	323	43*	Sell, 2004
323	324	45*	Sell, 2004
232	325	35*	Sell, 2004
325	326	40*	Sell, 2004
326	327	45*	Sell, 2004
87	328	35*	Sell, 2004
328	329	38*	Sell, 2004
329	330	41*	Sell, 2004
330	331	45*	Sell, 2004
328	332	38*	Sell, 2004
332	333	41*	Sell, 2004
333	334	45*	Sell, 2004
328	335	38*	Sell, 2004
335	336	41*	Sell, 2004
336	337	45*	Sell, 2004
87	338	40*	Alberts et al., 2002; Janeway et al., 2001
338	339	45*	Janeway et al., 2001
144	340	38*	Janeway et al., 2001
144	341	38*	Vickaryous and Hall, 2006
144	342	38*	Vickaryous and Hall, 2006
144	343	38*	Vickaryous and Hall, 2006
144	344	38*	Vickaryous and Hall, 2006
144	345	38*	Vickaryous and Hall, 2006
340	346	38*	Paxinos and Mai, 2004
340	347	38*	Paxinos and Mai, 2004
144	348	38*	Vickaryous and Hall, 2006
164	349	40*	Vickaryous and Hall, 2006
164	350	40*	Vickaryous and Hall, 2006
84	351	40*	Sell, 2004
351	352	45*	Sell, 2004
84	353	40*	Sell, 2004
353	354	45*	Sell, 2004
84	355	40*	Sell, 2004
355	356	45*	Sell, 2004
84	357	40*	Sell, 2004
357	358	45*	Sell, 2004, Towler and Gelberman, 2006
84	359	40*	Sell, 2004; Towler and Gelberman, 2006
359	360	45*	Sell, 2004
84	361	40*	Bianco et al., 2001; Sell, 2004
41	361	40*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
84	362	40*	Mochida, 2005
362	363	45*	Mochida, 2005
35	364	30*	Sadler, 2004
364	365	35*	Sadler, 2004
364	366	35*	Chen and Goldhamer, 2002
46	367	35*	Chen and Goldhamer, 2002; Kirschstein and Skirboll, 2001
365	368	40*	Vickaryous and Hall, 2006
365	369	40*	Vickaryous and Hall, 2006
365	370	40*	Vickaryous and Hall, 2006
365	371	40*	Sell, 2004
365	372	40*	Vickaryous and Hall, 2006
372	373	45*	Vickaryous and Hall, 2006

372	374	45*	Vickaryous and Hall, 2006
33	375	35*	Minasi et al., 2002
375	376	40*	Minasi et al., 2002
375	366	40*	Chen and Goldhamer, 2002
366	365	40*	Chen and Goldhamer, 2002
366	368	40*	Kirschstein and Skirboll, 2001
366	369	40*	Kirschstein and Skirboll, 2001
366	370	40*	Kirschstein and Skirboll, 2001
376	366	45*	Kirschstein and Skirboll, 2001
135	377	40*	Savage et al., 2003
135	378	40*	Savage et al., 2003
135	379	40*	Savage et al., 2003
135	380	40*	Savage et al., 2003
380	381	45*	Savage et al., 2003
380	382	45*	Savage et al., 2003
136	383	40*	Savage et al., 2003
160	384	45*	Freitas, 1999
160	385	45*	Freitas, 1999
139	386	45*	Vickaryous and Hall, 2006
58	387	35*	Sell, 2004
387	388	45*	Sell, 2004
387	389	45*	Sell, 2004
231	390	35*	Janeway et al., 2001
390	391	38*	Kirschstein and Skirboll, 2001
391	392	41*	Janeway et al., 2001
392	393	44*	Janeway et al., 2001
393	394	45*	Janeway et al., 2001
41	349	40*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
41	350	40*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
41	166	40	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
187	395	60	Sadler, 2004
146	396	60	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
146	397	60	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
72	398	60	Sadler, 2004
193	399	60	Sadler, 2004
97	400	63	Sadler, 2004
202	401	65	Sadler, 2004
202	402	65	Sadler, 2004
41	403	65	Nakashima and Redid, 2003; Sadler, 2004
70	404	65	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
97	405	65	Sadler, 2004
97	406	65	Paxinos and Mai, 2004; Sadler, 2004
156	407	75	Sadler, 2004
403	408	75	Nakashima and Redid, 2003; Sadler, 2004
188	409	60*	Sadler, 2004
409	410	75	Sadler, 2004
202	411	75	Sadler, 2004
402	412	75	Nakashima and Redid, 2003; Sadler, 2004
60	413	60*	Sadler, 2004
463	414	80	Freitas, 1999; Sadler, 2004
81	415	80	Paxinos and Mai, 2004; Sell, 2004
81	416	80	Paxinos and Mai, 2004; Sell, 2004
81	417	80	Paxinos and Mai, 2004; Sell, 2004
81	418	80	Paxinos and Mai, 2004; Sell, 2004
199	419	80*	Sadler, 2004
172	420	80	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
165	421	80	Sadler, 2004

98	422	95	Anglani et al., 2004; Sell, 2004
98	423	95	Anglani et al., 2004; Sell, 2004
99	424	95	Anglani et al., 2004; Sell, 2004
99	425	95	Anglani et al., 2004; Sell, 2004
99	426	95	Anglani et al., 2004; Sell, 2004
99	427	95	Anglani et al., 2004; Sell, 2004
99	428	95	Anglani et al., 2004; Sell, 2004
99	429	95	Anglani et al., 2004; Sell, 2004
99	430	95	Sell, 2004; Vickaryous and Hall, 2006
99	431	95	Sell, 2004; Vickaryous and Hall, 2006
100	432	95	Anglani et al., 2004; Sell, 2004
100	433	95	Anglani et al., 2004; Sell, 2004
101	434	95	Anglani et al., 2004; Sell, 2004
101	435	95	Anglani et al., 2004; Sell, 2004
101	436	95	Anglani et al., 2004; Sell, 2004
101	437	95	Anglani et al., 2004; Sell, 2004
102	438	95	Anglani et al., 2004; Sell, 2004
102	439	95	Anglani et al., 2004; Sell, 2004
79	440	95	Sadler, 2004
79	441	95	Sadler, 2004
185	442	100	Sadler, 2004
186	443	100	Sadler, 2004
79	444	100	Sadler, 2004
77	445	100	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
179	445	100	Paxinos and Mai, 2004
173	445	100	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
188	446	105	Sadler, 2004
146	447	108	Kirschstein and Skirboll, 2001; Paxinos and Mai, 2004; Temple, 2001
70	448	108	Paxinos and Mai, 2004; Sadler, 2004; Vickaryous and Hall, 2006
440	449	110	Sadler, 2004
79	450	110	Sadler, 2004
444	451	110	Sadler, 2004
419	452	110	Sadler, 2004
419	453	110	Sadler, 2004
419	454	110	Sadler, 2004
165	455	110	Sadler, 2004
441	456	110	Sadler, 2004
103	457	110	Sadler, 2004
447	458	115	Paxinos and Mai, 2004; Temple, 2001; Vickaryous and Hall, 2006
447	459	115	Paxinos and Mai, 2004; Temple, 2001; Vickaryous and Hall, 2006
403	460	115*	Nakashima and Redid, 2003; Sadler, 2004
159	461	90*	Paxinos and Mai, 2004
446	462	135*	Panteleyev et al., 2001
103	463	104	Sadler, 2004
140	464	60	Sadler, 2004
103	465	60	Sadler, 2004
188	466	75	Sadler, 2004
443	467	100*	Sadler, 2004
467	468	120*	Sadler, 2004
468	469	135*	Sadler, 2004
186	470	100*	Sadler, 2004
470	471	120*	Sadler, 2004
470	472	120*	Sadler, 2004
426	473	50*	Anglani et al., 2004
427	473	50*	Anglani et al., 2004
428	473	50*	Anglani et al., 2004

84	473	50*	Sell, 2004
101	474	60*	Lopez et al., 2001
184	475	60*	Freitas, 1999
475	476	75*	Freitas, 1999
165	477	80	Sadler, 2004
477	478	100*	Foster et al., 2002
477	479	100*	Foster et al., 2002
478	480	120*	Foster et al., 2002
480	481	130*	Foster et al., 2002
477	482	100*	Foster et al., 2002
414	483	100*	Freitas, 1999
465	484	100*	Freitas, 1999
465	485	100*	Freitas, 1999
185	486	100*	Sadler, 2004
450	487	125*	Freitas, 1999
450	488	125*	Freitas, 1999
451	489	125*	Vickaryous and Hall, 2006
451	490	125*	Vickaryous and Hall, 2006
449	491	125*	Vickaryous and Hall, 2006
449	492	125*	Vickaryous and Hall, 2006
140	493	100*	Freitas, 1999
493	494	125*	Freitas, 1999
209	495	60*	Otto, 2002
263	496	60*	Otto, 2002
265	496	60*	Otto, 2002
266	496	60*	Otto, 2002
267	496	60*	Otto, 2002
268	496	60*	Otto, 2002
269	496	60*	Otto, 2002
274	497	50*	Sell, 2004
276	498	50*	Sell, 2004
278	499	50*	Sell, 2004
394	500	50*	Janeway et al., 2001
500	501	55*	Janeway et al., 2001
500	502	55*	Janeway et al., 2001
500	503	55*	Vickaryous and Hall, 2006
500	504	55*	Vickaryous and Hall, 2006
500	505	55*	Vickaryous and Hall, 2006
500	506	55*	Vickaryous and Hall, 2006
318	507	50*	Janeway et al., 2001
507	508	55*	Janeway et al., 2001
508	509	75*	Janeway et al., 2001
508	510	75*	Janeway et al., 2001
508	511	75*	Janeway et al., 2001
508	512	75*	Janeway et al., 2001
508	513	75*	Janeway et al., 2001
507	514	60*	Janeway et al., 2001
514	515	75*	Janeway et al., 2001
515	516	90*	Janeway et al., 2001
324	517	50*	Sell, 2004
517	518	55*	Sell, 2004
331	519	60*	Sell, 2004
334	520	60*	Sell, 2004
337	521	60*	Sell, 2004
87	522	50*	Janeway et al., 2001
522	523	60*	Janeway et al., 2001
340	524	75*	Vickaryous and Hall, 2006

340	525	75*	Vickaryous and Hall, 2006
340	526	75*	Vickaryous and Hall, 2006
340	527	75*	Vickaryous and Hall, 2006
340	528	75*	Vickaryous and Hall, 2006
354	529	55*	Sell, 2004
529	530	65*	Sell, 2004
356	531	55*	Sell, 2004
41	531	55*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
359	532	55*	Sell, 2004
352	533	60*	Vickaryous and Hall, 2006
352	534	60*	Vickaryous and Hall, 2006
352	535	60*	Freitas, 1999; Vickaryous and Hall, 2006
41	536	80*	Santagati and Rijli, 2003; Sell, 2004
536	537	120*	Sell, 2004
41	538	100*	Santagati and Rijli, 2003
42	538	100*	Santagati and Rijli, 2003
41	539	75*	Santagati and Rijli, 2003
539	540	100*	Freitas, 1999
41	541	75*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
41	542	100*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
41	543	100*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
543	544	125*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
543	545	125*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
42	546	75*	Sadler, 2004
546	547	100*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
546	548	100*	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
64	549	60*	Sadler, 2004
549	550	100*	Vickaryous and Hall, 2006
540	551	100*	Vickaryous and Hall, 2006
188	552	120*	Sadler, 2004
552	553	135*	Vickaryous and Hall, 2006
103	554	75*	Sadler, 2004; Sell, 2004
554	555	100*	Vickaryous and Hall, 2006
188	556	120*	Potten and Booth, 2002
556	557	130*	Vickaryous and Hall, 2006
557	558	135*	Vickaryous and Hall, 2006
557	559	135*	Vickaryous and Hall, 2006
556	560	130*	Vickaryous and Hall, 2006
560	561	135*	Vickaryous and Hall, 2006
556	562	135*	Vickaryous and Hall, 2006
192	563	80*	Sadler, 2004
192	564	80*	Sadler, 2004
465	565	130*	Freitas, 1999
565	566	135*	Freitas, 1999
24	567	60*	Freitas, 1999
567	568	80*	Freitas, 1999
466	569	130*	Sadler, 2004
569	570	135*	Sadler, 2004
419	571	120*	Potten and Booth, 2002
571	572	135*	Potten and Booth, 2002; Vickaryous and Hall, 2006
64	573	100*	Sadler, 2004; Sell, 2004; Vickaryous and Hall, 2006
137	574	50*	Sadler, 2004
574	575	100*	Vickaryous and Hall, 2006
168	576	100*	Vickaryous and Hall, 2006
168	577	100*	Vickaryous and Hall, 2006
146	578	100*	Paxinos and Mai, 2004
70	579	50*	Sell, 2004

39	580	60*	Vickaryous and Hall, 2006
39	581	60*	Vickaryous and Hall, 2006
181	582	60*	Herrick and Mutsaers, 2004
582	583	80*	Herrick and Mutsaers, 2004; Sadler, 2004
71	584	75*	Sell, 2004
584	585	100*	Vickaryous and Hall, 2006
584	586	100*	Vickaryous and Hall, 2006
584	587	100*	Vickaryous and Hall, 2006
584	588	100*	Vickaryous and Hall, 2006
584	562	135*	Vickaryous and Hall, 2006
70	589	60*	Sell, 2004
161	590	75*	Sadler, 2004
590	591	100*	Sadler, 2004; Vickaryous and Hall, 2006
590	592	100*	Vickaryous and Hall, 2006
590	593	100*	Vickaryous and Hall, 2006
38	594	50*	Sadler, 2004
174	595	75*	Forge and Wright, 2002; Vickaryous and Hall, 2006
174	596	75*	Forge and Wright, 2002; Vickaryous and Hall, 2006
174	597	75*	Forge and Wright, 2002; Vickaryous and Hall, 2006
174	598	100*	Forge and Wright, 2002; Vickaryous and Hall, 2006
598	599	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
598	600	125*	Forge and Wright, 2002; Vickaryous and Hall, 2006
600	601	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
600	602	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
598	603	125*	Forge and Wright, 2002; Vickaryous and Hall, 2006
603	604	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
603	605	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
598	606	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
598	607	135*	Forge and Wright, 2002; Freitas, 1999
598	608	135*	Forge and Wright, 2002
174	609	100*	Forge and Wright, 2002; Sadler, 2004
609	610	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
609	611	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
405	612	100*	Forge and Wright, 2002; Vickaryous and Hall, 2006
612	613	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
612	614	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
405	615	100*	Forge and Wright, 2002; Vickaryous and Hall, 2006
405	616	100*	Vickaryous and Hall, 2006
405	617	100*	Vickaryous and Hall, 2006
405	618	100*	Forge and Wright, 2002; Vickaryous and Hall, 2006
618	619	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
618	620	135*	Forge and Wright, 2002; Vickaryous and Hall, 2006
191	621	100*	Freitas, 1999; Vickaryous and Hall, 2006
621	622	135*	Freitas, 1999; Vickaryous and Hall, 2006
621	623	135*	Freitas, 1999; Vickaryous and Hall, 2006
191	624	100*	Vickaryous and Hall, 2006
191	625	65	Vickaryous and Hall, 2006
97	625	65	Sadler, 2004
625	626	135*	Vickaryous and Hall, 2006
625	627	135*	Vickaryous and Hall, 2006
176	628	135*	Sadler, 2004
177	628	135*	Sadler, 2004
178	628	135*	Sadler, 2004
155	629	135*	Sadler, 2004
160	630	60*	Vickaryous and Hall, 2006
70	631	60*	Vickaryous and Hall, 2006
64	632	100*	Sadler, 2004

64	633	100*	Sadler, 2004
64	634	100*	Sadler, 2004
181	635	100*	Herrick and Mutsaers, 2004
181	636	100*	Herrick and Mutsaers, 2004
151	637	100*	Sell, 2004
211	638	120*	Sadler, 2004
45	498	50*	Vickaryous and Hall, 2006
415	639	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	640	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	641	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	642	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	643	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	644	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	645	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	646	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	647	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	648	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	649	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	650	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	651	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	652	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	653	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	654	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	655	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
415	656	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
416	657	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
416	658	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
416	659	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	660	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	661	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	662	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	663	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	664	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	665	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	666	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	667	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	668	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	669	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	670	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	671	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	672	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	673	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	674	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	675	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	676	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	677	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	678	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	679	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	680	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	681	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	682	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
417	683	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	684	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	685	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	686	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	687	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	688	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006

418	689	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	690	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	691	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
418	692	140	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
462	693	165	Panteleyev et al., 2001; Sadler, 2004
401	694	165	Sadler, 2004
460	695	165	Sadler, 2004
446	696	165	Nakashima and Redid, 2003; Sadler, 2004
170	697	170	Paxinos and Mai, 2004
170	698	185	Paxinos and Mai, 2004
170	699	190	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
170	700	205	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
77	700	205	Paxinos and Mai, 2004; Sadler, 2004
700	701	210	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
700	702	210	Paxinos and Mai, 2004; Sadler, 2004
700	703	210	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
461	704	220	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
466	705	225	Hennighausen and Robinson , 2005; Sadler, 2004
466	706	225	Hennighausen and Robinson , 2005; Sadler, 2004
446	707	240	Panteleyev et al., 2001; Sadler, 2004
446	708	240	Panteleyev et al., 2001; Sadler, 2004
446	709	240	Panteleyev et al., 2001; Sadler, 2004
209	710	140	Sadler, 2004
710	711	220*	Otto, 2002
710	712	200*	Otto, 2002
712	711	220*	Otto, 2002
81	713	140*	Sell, 2004
81	714	140*	Vickaryous and Hall, 2006
693	715	220*	Panteleyev et al., 2001
693	716	220*	Panteleyev et al., 2001
693	717	220*	Panteleyev et al., 2001
446	718	200*	Vickaryous and Hall, 2006
170	719	240*	Vickaryous and Hall, 2006
170	720	240*	Vickaryous and Hall, 2006
170	721	240*	Vickaryous and Hall, 2006
170	722	240*	Vickaryous and Hall, 2006
170	723	240*	Vickaryous and Hall, 2006
170	724	240*	Vickaryous and Hall, 2006
170	725	240*	Vickaryous and Hall, 2006
170	726	240*	Vickaryous and Hall, 2006
170	727	220*	Vickaryous and Hall, 2006
727	728	240*	Vickaryous and Hall, 2006
727	729	240*	Vickaryous and Hall, 2006
727	730	240*	Vickaryous and Hall, 2006
170	731	240*	Vickaryous and Hall, 2006
170	732	240*	Vickaryous and Hall, 2006
170	733	240*	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
170	734	240*	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
170	735	240*	Vickaryous and Hall, 2006
170	736	220*	Vickaryous and Hall, 2006
736	737	240*	Vickaryous and Hall, 2006
736	738	240*	Vickaryous and Hall, 2006
77	736	240*	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
77	739	240*	Paxinos and Mai, 2004; Vickaryous and Hall, 2006
77	740	240*	Vickaryous and Hall, 2006
171	741	240*	Vickaryous and Hall, 2006
171	742	240*	Vickaryous and Hall, 2006

171	743	240*	Vickaryous and Hall, 2006
179	744	240*	Vickaryous and Hall, 2006
179	745	240*	Vickaryous and Hall, 2006
179	746	240*	Vickaryous and Hall, 2006
173	747	240*	Vickaryous and Hall, 2006
173	748	240*	Vickaryous and Hall, 2006
173	749	240*	Vickaryous and Hall, 2006
579	750	240*	Vickaryous and Hall, 2006
579	751	240*	Vickaryous and Hall, 2006
579	752	240*	Vickaryous and Hall, 2006
706	753	225*	Hennighausen and Robinson , 2005
706	754	225*	Hennighausen and Robinson , 2005
705	755	200*	Hennighausen and Robinson , 2005
755	753	225*	Hennighausen and Robinson , 2005
755	754	225*	Hennighausen and Robinson , 2005
589	756	240*	Vickaryous and Hall, 2006
589	757	240*	Vickaryous and Hall, 2006
589	758	240*	Vickaryous and Hall, 2006
589	759	240*	Vickaryous and Hall, 2006
417	760	140*	Vickaryous and Hall, 2006
170	761	240*	Vickaryous and Hall, 2006
589	762	240*	Vickaryous and Hall, 2006
70	763	240*	Vickaryous and Hall, 2006
147	764	240*	Vickaryous and Hall, 2006
147	765	240*	Vickaryous and Hall, 2006
147	766	240*	Vickaryous and Hall, 2006
147	767	240*	Vickaryous and Hall, 2006
147	768	240*	Vickaryous and Hall, 2006
147	769	240*	Vickaryous and Hall, 2006
147	770	240*	Vickaryous and Hall, 2006
147	736	220*	Vickaryous and Hall, 2006
448	764	240*	Vickaryous and Hall, 2006
448	765	240*	Vickaryous and Hall, 2006
448	766	240*	Vickaryous and Hall, 2006
448	767	240*	Vickaryous and Hall, 2006
448	768	240*	Vickaryous and Hall, 2006
448	769	240*	Vickaryous and Hall, 2006
448	770	240*	Vickaryous and Hall, 2006
448	740	240*	Vickaryous and Hall, 2006
404	771	240*	Vickaryous and Hall, 2006
631	772	240*	Vickaryous and Hall, 2006
631	773	240*	Vickaryous and Hall, 2006
634	774	160*	Sadler, 2004
774	775	200*	Sadler, 2004
775	776	240*	Sadler, 2004
775	777	240*	Sadler, 2004
634	778	200*	Sadler, 2004
778	779	240*	Sadler, 2004
464	780	200*	Sadler, 2004
421	781	140*	Vickaryous and Hall, 2006
455	782	140*	Sell, 2004
407	783		Sadler, 2004
407	784		Sadler, 2004
783	784		Sadler, 2004
783	785		Sadler, 2004
52	786		Sadler, 2004
786	787		Sadler, 2004

787	788	Sadler, 2004
788	789	Sadler, 2004
789	790	Alberts et al., 2002
790	791	Alberts et al., 2002
791	792	Alberts et al., 2002
215	793	Vickaryous and Hall, 2006
216	793	Vickaryous and Hall, 2006
105	793	Vickaryous and Hall, 2006
104	793	Vickaryous and Hall, 2006
106	793	Vickaryous and Hall, 2006
215	794	Vickaryous and Hall, 2006
216	794	Vickaryous and Hall, 2006
215	795	Vickaryous and Hall, 2006
216	795	Vickaryous and Hall, 2006
105	795	Vickaryous and Hall, 2006
215	796	Vickaryous and Hall, 2006
216	796	Vickaryous and Hall, 2006
215	797	Vickaryous and Hall, 2006
216	797	Vickaryous and Hall, 2006
105	798	Vickaryous and Hall, 2006
105	799	Vickaryous and Hall, 2006
105	800	Vickaryous and Hall, 2006
106	800	Vickaryous and Hall, 2006
104	800	Vickaryous and Hall, 2006
105	801	Vickaryous and Hall, 2006
105	802	Vickaryous and Hall, 2006
106	802	Vickaryous and Hall, 2006
104	802	Vickaryous and Hall, 2006
105	803	Vickaryous and Hall, 2006
105	804	Vickaryous and Hall, 2006
41	805	Santagati and Rijli, 2003
41	806	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
41	807	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
556	810	Sadler, 2004
41	811	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
41	812	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
812	808	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
812	809	Santagati and Rijli, 2003; Vickaryous and Hall, 2006
707	813	Panteleyev et al., 2001
707	814	Panteleyev et al., 2001
707	815	Panteleyev et al., 2001
97	816	Freitas, 1999
816	817	Vickaryous and Hall, 2006
560	818	Stoeckelhuber et al., 2003
818	819	Freitas, 1999; Vickaryous and Hall, 2006
74	820	Freitas, 1999
820	821	Vickaryous and Hall, 2006
590	822	Freitas, 1999
822	823	Vickaryous and Hall, 2006
822	824	Vickaryous and Hall, 2006
84	825	Sell, 2004
825	826	Vickaryous and Hall, 2006
825	827	Vickaryous and Hall, 2006
825	828	Vickaryous and Hall, 2006
825	829	Vickaryous and Hall, 2006
41	830	Szeder et al., 2003
830	831	Szeder et al., 2003; Vickaryous and Hall, 2006

419	830	Vickaryous and Hall, 2006
105	832	Freitas, 1999; Vickaryous and Hall, 2006
104	832	Vickaryous and Hall, 2006
106	832	Vickaryous and Hall, 2006
105	833	Vickaryous and Hall, 2006
105	835	Vickaryous and Hall, 2006
496	834	Freitas, 1999; Vickaryous and Hall, 2006
496	832	Vickaryous and Hall, 2006
143	836	Vickaryous and Hall, 2006
836	837	Vickaryous and Hall, 2006
836	838	Vickaryous and Hall, 2006
17	838	Vickaryous and Hall, 2006
402	839	Nakashima and Redid, 2003
839	840	Nakashima and Redid, 2003; Vickaryous and Hall, 2006
340	841	Vickaryous and Hall, 2006
340	842	Vickaryous and Hall, 2006
140	843	Freitas, 1999
843	844	Freitas, 1999; Vickaryous and Hall, 2006
140	845	Freitas, 1999
845	846	Freitas, 1999; Vickaryous and Hall, 2006
56	847	Freitas, 1999; Sadler, 2004; Sell, 2004
847	848	Freitas, 1999; Vickaryous and Hall, 2006
365	849	Vickaryous and Hall, 2006
28	849	Vickaryous and Hall, 2006
233	849	Sadler, 2004
135	850	Vickaryous and Hall, 2006
405	851	Sadler, 2004; Vickaryous and Hall, 2006
174	852	Vickaryous and Hall, 2006
174	853	Vickaryous and Hall, 2006
174	854	Vickaryous and Hall, 2006
84	855	Vickaryous and Hall, 2006
33	856	Vickaryous and Hall, 2006
64	857	Vickaryous and Hall, 2006
857	858	Vickaryous and Hall, 2006
840	859	Freitas, 1999
408	860	Freitas, 1999
194	861	Sadler, 2004
194	862	Sadler, 2004
61	863	Sadler, 2004
61	864	Sadler, 2004
61	865	Sadler, 2004
61	866	Sadler, 2004
866	867	Sadler, 2004
866	868	Sadler, 2004
866	869	Sadler, 2004
866	870	Sadler, 2004
866	871	Sadler, 2004
866	872	Sadler, 2004
495	270	Otto, 2002
224	361	Bianco et al., 2001
225	361	Bianco et al., 2001
226	361	Bianco et al., 2001
227	361	Bianco et al., 2001
228	361	Bianco et al., 2001
229	361	Bianco et al., 2001
230	361	Bianco et al., 2001
361	349	Sell, 2004

361	350		Sell, 2004
361	166		Sell, 2004
361	533		Sell, 2004
361	534		Sell, 2004
361	473		Vickaryous and Hall, 2006
531	473		Sell, 2004
531	529		Alberts et al., 2002
531	530		Alberts et al., 2002
531	150		Alberts et al., 2002
531	164		Alberts et al., 2002
531	349		Alberts et al., 2002
531	350		Alberts et al., 2002
531	166		Alberts et al., 2002
531	533		Alberts et al., 2002
531	534		Alberts et al., 2002
460	408		Nakashima and Redid, 2003
136	873	50*	Savage et al., 2003

CHAPTER 6

A multi-agent simulation for the enzyme kinetics of invertase

Abstract

In this work we have developed a three-dimensional multi-agent-based model to investigate the enzyme kinetics of invertase. The model has been validated with experiments performed at different temperatures. For each concentration, the sucrose cleavage was measured by the absorbance at different time. In the computational model, we have five types of agents: invertase, sucrose, buffer, glucose, and fructose. There are empty spaces to simulate the mobility of the molecules. These obey a Maxwell-Boltzmann velocity distribution and move by using the Moore neighborhood. In our rules, buffer does not change its state, invertase can denature, and sucrose can change to glucose and fructose. The results of our computational model reproduce the Michaelis-Menten kinetics of reducing sugar formation. Additionally, they show that the velocity of cleavage is related to the initial fraction of sucrose, initial fraction of invertase, and depend on the temperature which the reaction takes places.

Keywords: computational model, Maxwell-Boltzmann velocity distribution, enzymatic reaction, sucrose cleavage, invertase.

6.1. Introduction

Enzymes are biological catalysts that increase the rates of chemical reactions. A classic enzyme has an active site in which a substrate can bind. Once situated in the active site, the substrate can become more reactive through conformational changes or polarizations produced by intermolecular interaction with the enzyme. Also, the enzyme can promote the substrate division into one or more products. In the experiments, there are many physical-chemistry parameters that can affect the enzyme performance. The principal parameters are substrate concentration, enzyme

concentration, temperature, pH and molecular compounds of the solution (Koolman and Roehm, 2005).

For instance, in a well-know degradation reaction, the invertase enzyme (β-D-fructofuranoside fructohydrolase, EC 3.2.1.26) can catalyze the sucrose (α-D-glucopyranosyl-(1,2)-β-D-fructofuranoside) cleavage (Reid and Abratt, 2005; Sturm and Tang, 1999). Invertase is a glycoside hydrolase that catalyze the irreversible cleavage of sucrose into two monosaccharides, namely, glucose and fructose. This enzyme is very important for biotechnological applications, such as the production of non-crystallizable sugars, mostly used for the preparation of jams, candies and soft-centered chocolates (Rubio et al., 2002). Invertase has significant advantages in the manufacture of syrup due to the absence of byproduct formation (Almeida et al., 2005). Recent research suggests the potential use of invertase in the production of lactic acid obtained by the fermentation of cane molasses in ethanol production (Kirk and Doelle, 1993; Lee and Huang, 2000). Also, invertase can be used in biosensors for sucrose determination (Park et al., 2001).

The action of invertase occurs through the hydrolyzis of terminal non-reducing β-D-fructofuranoside residues in β-D-fructofuranosides (Heil et al., 2005; Sturm and Tang, 1999). In the end of the reaction, the number of molecules of glucose and fructose is the same. However, this exact quantity is not directly measured. This number is visualized indirectly through the quantity of reducing sugar (Koolman and Roehm, 2005). The native enzyme is maintained by a delicate balance of non-covalent forces and hydrogen bonds. The increase of temperature can induce disruption of these interactions and the enzyme unfolding. The degree of enzyme unfolding can be observed by different techniques and the denaturation is irreversible for the majority of enzymes (Gomes et al., 2007).

Enzyme reactions are dynamic, but the information about the experiments is not dynamic due to the time interval separating observations. The limitation would be improved if all concentrations were detected at small time intervals (Koolman and Roehm, 2005). An approach to this limitation is the creation of computational models in which the concentrations are known in all time. In this direction, different theoretical models have been proposed to investigate the enzymatic reaction. Examples include a mathematical model for the inulinase (Santos et al., 2007), as well as models based on cellular automata for the solute in a solvent (Kier and Cheng, 1994), chemical kinetics (Seybold et al., 1997), enzyme kinetics (Kier and Cheng, 2000; Kier et al., 1996; Weimar, 2002), biochemical phenomena (Kier et al., 1999), and iteration among molecules (Kier, 2000). Usually, these models do not include the Maxwell-Boltzman velocity distribution for the molecules. In this way, we have developed a three-dimensional multi-agent-based model for the reaction of invertase with sucrose by using this velocity distribution. Also, in order to have a control of the

parameters affecting the enzyme kinetics and validate our computational model, we have performed a set of experiments for sucrose degradation.

6.2. Material and methods

6.2.1. Chemicals

Sephadex G-25, bovine serum albumin, and 3,5-dinitrosalicylicacid were purchased from Sigma Chemical Co. (St Louis, MO, USA). All the other chemicals used were of high-quality analytical grade.

6.2.2. Enzyme

For extraction of invertase, 1 g dry yeast cells (*Saccharomyces cerevisae*) was provided by Pak Gida (Istanbul - Turkey) and macerated in 30 mL of ethyl eter and silica. The cells were kept for 30 min at room temperature and centrifuged on 10 g x 10 min. The enzyme purification was done through the application of supernatant on Sephadex G-25 column (1 x 5 cm) washed with 50 mM acetate buffer (pH 4.5). The elution was done with same buffer with flow rate of 2 mL/min. The fractions containing invertase were pooled for subsequent assay of activity and protein determination. The unused fractions were kept at 4 °C.

6.2.3. Enzyme activity

The reaction was initiated by adding enzyme in the solution. Sucrose hydrolysis was catalyzed by invertase in sodium acetate buffer (50 mM, pH = 4.5). Invertase activity was measured according to the concentration of reducing sugars released from sucrose. Reducing sugars were measured according to the DNS acid method (Miller, 1959). One unit of invertase activity is defined as the amount of enzyme that catalyzes the hydrolysis of 1.0 μmol reducing sugar per minute, under the conditions of assay. All fractions were assayed for absorbance at 540 nm.

6.2.4. Protein determination

Protein concentration was measured by the Bradford method (Bradford, 1976). For the calibration curve it was used bovine serum albumin (Sigma Chemical Co., St. Louis, MO, USA) as a standard.

6.2.5. Temperature and pH

The optimum temperature for invertase was tested under standard assay conditions at several temperatures. The temperatures (30, 40, 50, 60, and 70 °C) were controlled by circulating water bath. Samples without invertase were made and the invertase activity was determined with the fix value of pH 4.5.

6.2.6. Kinetics parameters of invertase

The effect of substrate concentration was studied by determination of the initial velocities (V_o) at different concentrations (0.0025; 0.005; 0.0075; 0.01; 0.015; 0.02; 0.025; 0.03 mM). The Michaelis–Menten constant (Michaelis and Menten, 1913) (k_M) and the maximum velocity (V_{max}) were determined using the Lineweaver-Burk method (Lineweaver and Burk, 1934).

6.2.7. Statistical Analyses

Average values of triplicates (differing about 5%) were calculated. The data obtained from the study were analyzed using linear or quadratic regression.

6.2.8. Computational model

The description of our computational model follows the standard Overview, Design concepts, and Details (ODD) protocol for individual-based and agent-based models (Grimm et al., 2006).

6.2.8.1. Purpose

The aim of this computational model is to study and explore the possibility of an agent-based model to simulate the sucrose hydrolysis, *in silico*, giving the exact concentration of the formed products.

6.2.8.2. State variables and scales

The three-dimensional regular lattice consists of a grid with 100 x 100 x 100 sites to represent the solution. Each site state is selected and distributed on a regular lattice. This lattice

possesses periodic boundary conditions and each site lattice corresponds to the space region that can be an empty space or can be occupied by a single type of autonomous agent. Our model has different types of agents to simulate the reaction of invertase with sucrose. Also, each type of agent represents a different type of molecule. The agents correspond to molecular groups with a similar pattern of behavior because the number of molecules in a solution is large. Time and space are discrete and all actions occur at constant intervals called time steps.

This model includes five different types of agents: invertase, sucrose, buffer, glucose and fructose. The only parameter is the total number of agents. The quantity of invertase and buffer is constant during the simulation. The quantity of buffer is given by the difference between the total number of agents and the quantity of sucrose and invertase. In our model, 100 time steps correspond to the experimental data of 50 minutes; therefore, each time step corresponds to approximately 30 seconds of the experiment.

6.2.8.3. Process overview and scheduling

All types of agents have an initial random distribution and can move through the empty spaces. Each type of agent is chosen randomly and moved to empty space. These types of agents have a Maxwell-Boltzmann velocity distribution and move by using the three-dimensional Moore neighborhood (Wolfram, 1986). In this type of neighborhood each site has 26 neighbors sites. This type of velocity distribution is important because the temperature and denaturation effect can be included. Additionally, there are groups of molecules with different velocities. Thus, we use a Maxwell-Boltzmann velocity distribution as defined below:

$$N_v = 4\pi N\left(\frac{1}{2\pi T}\right)^{3/2} v^2 e^{-v^2/2T} \qquad (1)$$

Eq. (1) was simplified to represent N molecules with unit mass and the Boltzmann constant set to 1.

6.2.8.4. Design concepts

Emergence: Population kinetics emerges from the behavior of different types of molecules.
Sensing: Agents know their own state, recognize the state of the other agents and apply the transition rules according to the state of their first neighbors.

Interaction: The system evolution is determined by local interactions between the invertase, sucrose, buffer, glucose, and fructose.
Observation: The total number of glucose and fructose are investigated at each time step.

6.2.8.5. Initialization

The spatial distribution pattern of invertase, sucrose and buffer is random. Initially, the solution does not possess glucose and fructose; these two types of agents only appear during the time evolution. Also, there are not enzymes denatured in the beginning of the simulation.

6.2.8.6. Input

In the experiment, we know the temperature and the initial fractions of invertase, sucrose and buffer. Therefore, in our computational model the temperature and the initial fractions of invertase, sucrose and buffer are given by the experimental data.

6.2.8.7. Submodels

The time evolution is equally run in the complete lattice and it is obtained when the transition rules are applied to all agents randomly selected. Each site changes its state according to a local rule, which depends only on the adjacent neighbors. The transition rules are uniformly applied to each molecule and they depend on a set of rules that establish the state of all sites. They control the joining and breaking of sucrose by invertase. Our model has one rule to simulate the reaction of invertase with sucrose and one rule to simulate the enzyme denaturation. A schematic representation to explain the rules for sucrose cleavage and enzyme denaturation is shown in Fig. 1. The transitions rules are described as follow:

(i) If there is a sucrose in the central site; if there is an empty space in the left of the sucrose, and if the site up or down of the sucrose is an invertase; the sucrose site changes into glucose and the empty space site changes into a fructose (Koolman and Roehm, 2005). This rule is specific due to the lock and key model of enzyme-substrate interaction.

(ii) In the neighborhood of an invertase; if there is a molecule with the velocity greater than a maximum velocity at 50^0 C, the invertase changes its configuration, i.e., the invertase becomes

denatured (Koolman and Roehm, 2005). This occurs because the structure of most enzymes can unfold by increase of temperature, disabling the enzyme.

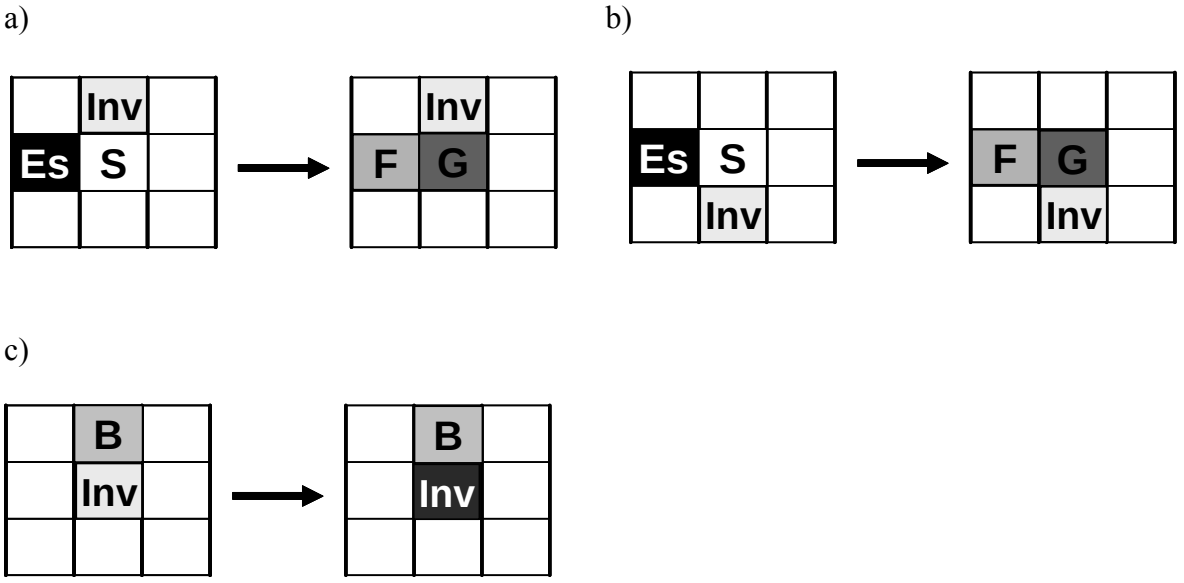


Fig. 1. Schematic representation of the transition rules for our enzyme kinetics model. The Moore neighborhood is represented in two-dimensions for clarity. However, we use the three-dimensional Moore neighborhood. a) If there is a sucrose (S) in the central site; if there is an empty space (Es) in the left of the S, and if the site up of the S is an invertase (Inv); the S site changes into glucose (G) and the Es site changes into a fructose (F) (Koolman and Roehm, 2005). b) If there is a S in the central site; if there is an Es in the left of the S, and if the site down of the S is an Inv; the S site changes into G and the Es site changes into a F (Koolman and Roehm, 2005). c) In the neighborhood of an Inv; if there is a molecule with the velocity greater than a maximum velocity at 50^0 C, the Inv changes its configuration (Koolman and Roehm, 2005).

6.3. Results and Discussion

For each set of parameters, the average value of 20 simulation runs was taken in order to better describe the reducing sugar formation. The parameters optimization was done by using the fit to our experimental results. In the experiment the total volume of solution is 0.75 ml. The quantity of buffer is given by the difference between the total volume of solution and the quantity of the others types of molecules (invertase and sucrose). In all experiments, the quantity of sucrose is 0.25 ml, i.e., the quantity of sucrose corresponds to 33.33% of the solution. However, we can vary the molarity of sucrose. The larger molarity of sucrose that we measure was 0.03 M. Therefore, in our computational data, the 0.03 M sucrose corresponds to fraction of sucrose 0.333, the 0.02 M sucrose corresponds to fraction 0.222 and so forth.

The data obtained by computational model were compared with the experimental results. The relation found between experimental absorbance (A) and the number of glucose or fructose agents ($N_{g,f}$) is given by $A = \alpha N_{g,f}$, where α represents the dimensional parameter. To determine

the α parameter, the experimental value of the absorbance at 50 minutes was divided by the final quantity of invertase ($N_{g,f}$ at iteration 100). Hence, we have three distinct values for the three curves with different quantity of invertase. In Figure 2 we show these values and the curve has a linear behavior. The curve is given by:

$$\alpha = \beta I + \gamma \qquad (2)$$

where I is the quantity of invertase (ml), β is related to the fraction of invertase affecting the absorbance, and γ is related to the fraction of reducing sugar formation affecting the absorbance. The values of β and γ are 7.8 x 10^{-5} and 2.4 x 10^{-6}, respectively, with $R^2 = 0.999$. This equation only depends on the quantity of enzyme in the solution, i.e., it is independent of the substrate concentration. In all curves, we use the equation 2 to multiply our simulated data and compare with the experimental results.

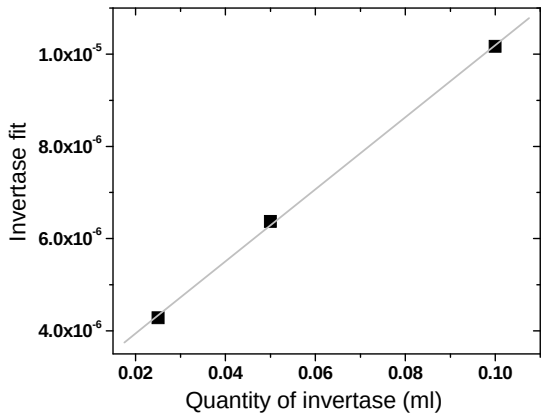


Fig. 2. Linear curve to obtain the α constant.

The absorbance for different time and fractions of enzyme and substrate is shown in Figure 3. We can observe that the quantity of reducing sugar measured tends to stabilize during the time evolution. In our experimental results, the occurrence of sucrose cleavage is more pronounced in the first minutes of the enzymatic reaction. Also, the larger the concentration of enzyme and substrate increases, the larger the first absorbance measure (0.25 min). However, our computational results start on zero because initially the formation of glucose and fructose is not present. Therefore, the larger the time, the better the concordance between computational and experimental data.

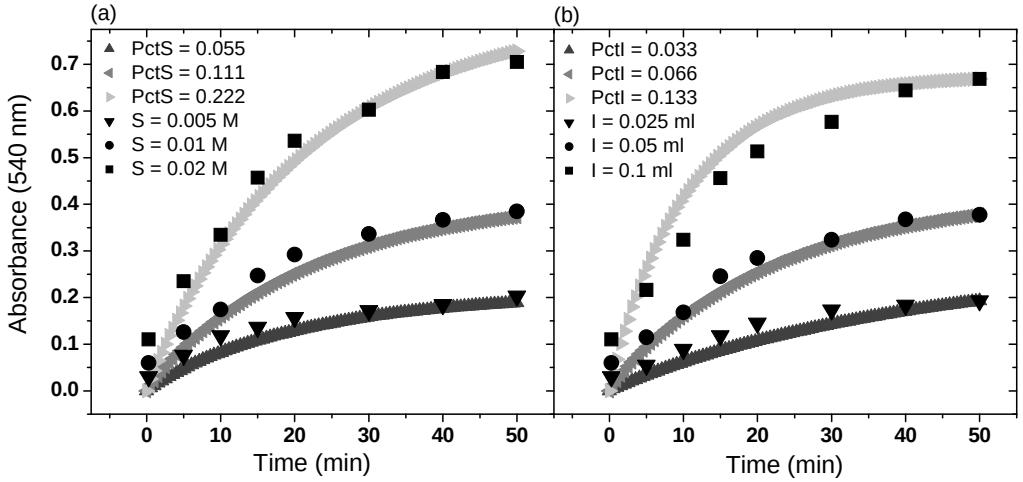


Fig. 3. Absorbance as a function of time for different fractions of sucrose (S) and invertase (I). In the experimental and computational data the temperature (T) is 50 °C and total volume of solution (VS) is 0.75 ml. In the computational data the fraction of lattice-site occupied (So) is 0.6. a) Substrate stability. In this experiment we use I = 0.05 ml; therefore, in the simulation we use PctI = 0.0666. b) Enzyme stability. In this experiment we use 0.01 M sucrose; therefore, in the simulation we use PctS = 0.111.

The plot in Figure 4 shows the Lineweaver-Burk representation of the experimental and computational results. From this curve, we have determined the kinetics constants $v_{\max}$ and k_M. The experimental $v_{\max}$ and k_M are 0.66 μmols.min^{-1} and 0.04 mM, respectively. For comparison, the computational $v_{\max}$ and k_M are 0.79 μmols.min^{-1} and 0.07 mM, respectively. The equation $y = 0.02x + 1.52$ with $R^2 = 0.999$ was estimate for the experimental data and $y = 0.09x + 1.26$ with $R^2 = 0.999$ for the computational data. The ratio $k_M / v_{\max}$ corresponds to the catalytic velocity constant. The experimental k_{catal} is 0.06 min^{-1} and the computational k_{catal} is 0.089 min^{-1}.

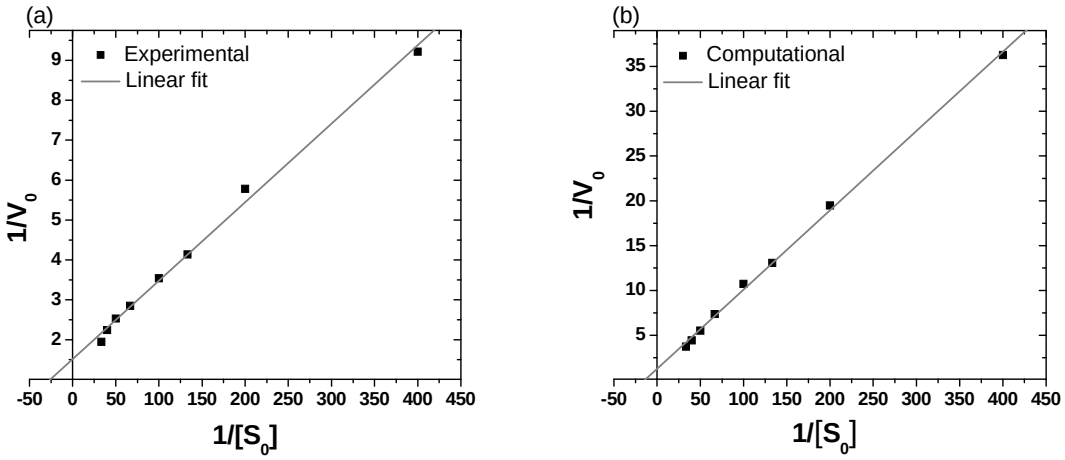


Fig. 4. Lineweaver-Burk representation of substrate concentration effect on the activity of invertase from *Saccharomyces cerevisae*. The initial velocity (v_o) is given by μmols.min^{-1}. The others parameters used were I = 0.05 ml, T = 50 °C, and VS = 0.75. a) Experimental data. b) Computational data. In this simulation we use So = 0.6. (See Fig. 3 for the meaning of the labels).

Figure 5 shows the kinetics of enzyme denaturation as a function of time for different temperatures. Our computational data show that initially the quantity of denatured invertase per time unit is large and in the end of the simulation this quantity is small. This is due to the reduction in the number of invertase that can denature during the time evolution. Also, the larger the temperature, the larger the quantity of denatured invertase. The denaturation process occurs more rapidly when the temperature increases because the velocity of the molecules depends on the temperature.

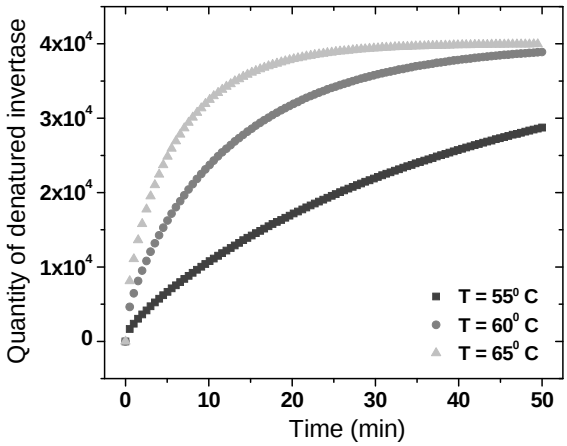


Fig. 5. Kinetics of enzyme denaturation as a function of time for different temperatures. The others parameters used were I = 0.05 ml, 0.02 M sucrose, So = 0.6, and VS = 0.75. (See Fig. 3 for the meaning of the labels).

The absorbance as a function of temperature for different fractions of lattice-site occupation is shown in Figure 6. To plot this graph, we have obtained the absorbance at 15 minutes of reaction. The enzyme exhibited maximum activity at 50 °C and it becomes denatured when the temperature is superior to 50 °C in both experimental and computational data. Therefore, the absorbance of reducing sugar increases up to 50 °C and after this temperature the absorbance decreases due to the invertase denaturation.

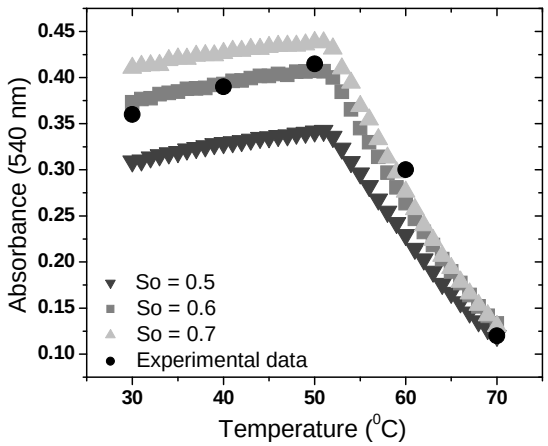


Fig. 6. Absorbance as a function of temperature for different fractions of lattice-site occupation. The others parameters used were I = 0.05 ml, 0.02 M sucrose, and VS = 0.75. (See Fig. 3 for the meaning of the labels).

6.4. Conclusion

In this paper, we have presented a multi-agent-based model to investigate the sucrose cleavage by invertase. We demonstrate here that this model, generated from appropriate rules, may reproduce some of the useful characteristics of enzyme kinetics. Also, it gives the possibility to know the concentrations of all molecules at each time step. As far as we know this paper presents the first computational model for enzymatic reaction based on regular lattice that includes a three-dimensional lattice, effect of temperature and denaturation. The inclusion of temperature and denaturation in the model was only possible by using the Maxwell-Boltzmann velocity distribution.

The computational and experimental data were compared for the evolution of reducing sugar formation, initial velocity and temperature effect, giving a good agreement. Our results reproduce the Michaelis-Menten kinetics and give a good fitting using the Lineweaver-Burk method. They show that the invertase from *Saccharomyces cerevisae* has maximum activity at 50 °C and the denaturation starts after this temperature. Finally, we have obtained that the temperature and the initial fraction of invertase sucrose affect the reducing sugar formation.

Acknowledgements

This work was supported by FAPESB and CNPq.

References

Almeida, A.C.S., Araújo, L.C., Costa, A.M., Abreu, C.A.M., Lima, M.A.G.A., Palha, M.A.P.F. (2005) Sucrose hydrolysis catalyzed by auto immobilized invertase into intact cells of *Cladosporium cladosporioides*. Eletron. J. of Biotechn., 8, 54-62.

Bradford, M.M. (1976) A rapid and sensitive method for the quantization of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem., 72, 248-254.

Gomes, E., Guez, M.A.U., Martin, N., Silva, R. (2007) Enzimas termoestáveis: fontes, produção e aplicação industrial. Quim. Nova, 30, 136-145.

Grimm, V., et al. (2006) A standard protocol for describing individual-based and agent-based models. Ecol. Model., 198, 115-126.

Heil, M., Buchler, R., Boland, W. (2005) Quantification of invertase activity in ants under field conditions. J Chem Ecol., 31, 431-437.

Kier, L.B. (2000) A cellular automata model of bond interactions among molecules. J. Chem. Inf. Comput. Sci., 40, 1285-1288.

Kier, L.B., Cheng, C.-K. (1994) A cellular automata model of an aqueous solution. J. Chem. Inf. Comput. Sci., 34, 1334-1337

Kier, L.B., Cheng, C.-K. (2000) A cellular automata model of an anticipatory system. J. Mol. Graph. Model., 18, 29–32.

Kier, L.B., Cheng, C.-K., Testa, B. (1999) Cellular automata models of biochemical phenomena. Fut. Generat. Comput. Syst., 16, 273–289.

Kier, L.B., Cheng, C.-K., Testa, B., Carrupt, P.-A. (1996) A cellular automata model of enzyme kinetics. J. Mol. Graph. Model., 14, 227-231.

Kirk, L.A., Doelle, H.W. (1993) Rapid ethanol production from sucrose without byproduct formation. Biotechnol. Lett., 15, 985-990.

Koolman, J., Roehm, K.-H. (2005) Color atlas of biochemistry. Second ed., Thieme, Stuttgart, New York.

Lee, W.C., Huang, C.T. (2000) Modelling of ethanol production using *Zymomonas mobilis* ATCC 10988 grown on the media containing glucose and fructose. Biochem. Eng. J., 4, 217-227.

Lineweaver, H., Burk, D. (1934) The determination of enzyme dissociation constants. J. Am. Chem. Soc., 56, 658–666.

Michaelis, L., Menten, M.L. (1913) Die kinetik der invertinwirkung. Biochem. Z, 49, 333-369.

Miller, G. (1959) Use of dinitrosalicylic acid reagent for detection of reducing sugar. Anal. Chem., 31, 426-428.

Park, J.K., Ro, H.S., Kim, H.S. (1991) A new biosensor for specific determination of sucrose using an oxidoreductase of Zymomonas mobilis and invertase. Biotechnol. Bioeng., 38, 217-223.

Reid, S. J., Abratt, V.R. (2005) Sucrose utilization in bacteria: genetic organisation and regulation. Appl. Microbiol. Biotechnol., 67, 312–321.

Sturm, A., Tang, G.Q. (1999) The sucrose-cleaving enzymes of plants are crucial for development, growth and carbon partitioning. Trends Plant. Sci., 4, 401–407.

Rubio, M.C., Runco, R., Navarro, A.R. (2002) Invertase from a strain of *Rhodotorula glutinis*. Phytochemistry, 61, 605-609.

Santos, A.M.P., Oliveira, M.G., Maugeri, F. (2007) Modelling thermal stability and activity of free and immobilized enzymes as a novel tool for enzyme reactor design. Bioresource Technol., 98, 3142–3148.

Seybold, P.G., Kier, L.B., Cheng, C.-K. (1997) Simulation of first-order chemical kinetics using cellular automata. J. Chem. Inf. Comput. Sci., 37, 386-391

Weimar, J. R. (2002) Cellular automata approaches to enzymatic reaction networks. Lect. Notes Comput. Sci., 2493, 294-303.

Wolfram, S. (1986) Theory and applications of cellular automata. World Scientific, Singapore.

Conclusão geral

Nesta tese foram desenvolvidos cinco modelos computacionais para três sistemas biológicos diferentes. Os sistemas estudados foram à rede de diferenciação celular, a doença de Chagas e cinética enzimática da invertase. Nossos resultados para a rede de diferenciação celular permitem uma visão integrada das propriedades dinâmicas e topológicas do processo de diferenciação revelando características essenciais no desenvolvimento embrionário humano. A rede construída evidencia a sensibilidade de algumas células, pois mutações em células muito conectadas podem alterar de forma substancial o mecanismo de diferenciação celular. Na modelagem da doença de Chagas após o transplante de células-tronco podemos concluir que o padrão de concentração é o fator mais importante para caracterizar a cinética de regeneração do tecido chagásico após o transplante de células-tronco. Na evolução da doença de Chagas fomos capazes de entender a participação dos diferentes tipos de células no desenvolvimento da cardiomiopátia chagásica crônica. A inclusão da distribuição de velocidade de Maxwell-Boltzamann, no trabalho sobre a cinética enzimática da invertase, nos permitiu modelar de forma eficiente o efeito da temperatura e a emergência do padrão de desnaturação. Nos modelos sobre doença de Chagas e cinética enzimática conseguimos obter bons ajustes entre os dados computacionais e os dados experimentais. Estes resultados, juntamente com os resultados da rede de diferenciação celular, nos permitiram uma visão mais ampla dos mecanismos microscópicos que determinam os padrões macroscópicos emergentes destes sistemas complexos.