

BACHELOR-ARBEIT

Evolution vielfaltserzeugender Mechanismen im Immunsystem

-

Evolution of diversity generating mechanisms in the immune system

Bachelorstudiengang Molecular Life Science



UNIVERSITÄT ZU LÜBECK

von

Meike Paschen

Gutachter:

1. Prof. Dr. Till Tantau

Institut für Theoretische Informatik
Universität zu Lübeck

2. Dr. Steffen Möller

Institut für Neuro- und Bioinformatik
Universität zu Lübeck

Datum:

Erklärung:

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Lübeck, den 12.8.2009

Meike Paschen

Inhaltsverzeichnis:

1. Zusammenfassung.....	1
2. Einführung.....	2
3 Grundlagen der Vielfalterzeugung.....	5
3.1 Somatische Rekombination.....	5
3.2 Genkonversion.....	7
4. Simulation.....	8
4.1 Genetischer Algorithmus.....	8
4.2 Modelle der vielfaltserzeugenden Mechanismen.....	9
4.2.1 Modell der Somatischen Rekombination.....	9
4.2.2 Modell der Genkonversion	10
4.2.3 Modell der Antikörpergenerierung.....	11
4.2.4 Repräsentation der Antigene.....	12
4.3 Fitnessberechnung.....	13
4.4 Generierung einer neuen Population.....	15
4.5 Implementierung.....	16
5. Ergebnisse	17
5.1 Versuche auf Basis von Standardbedingungen	17
5.1.1 Fitnessentwicklung bei Genkonversion und Somatischer Rekombination	17
5.1.2 Einfluss der Populationsgröße.....	18
5.1.3 Einfluss der Antikörperanzahl.....	19
5.1.4 Einfluss der Antigenanzahl	20
5.1.5 Einfluss der Antigenlänge.....	21
5.1.6 Einfluss der Anzahl an Genkonversionsevents.....	22
5.1.7 Einfluss der Generierung einer neuen Antigenbibliothek.....	23
5.2. Versuche auf Basis von Simulationserweiterungen.....	24
5.2.1 Einfluss der kontinuierlichen Erzeugung neuer Antigene.....	24
5.2.2 Entwicklung der Genvielfalt über Generationen.....	26
6. Diskussion.....	27
7. Literaturverzeichnis.....	29
8. Anhang.....	30

1. Zusammenfassung

Das Immunsystem höherer Organismen ist ein sehr komplexes System, welches ca. 10^{11} unterschiedliche Antikörper [Janeway et. al., 2005, 6, S. 136] erzeugt, obwohl die DNA mit einer Länge von $3 \cdot 10^9$ Basenpaaren für viel weniger kodiert. Somatische Rekombination und Genkonversion sind zwei von einander unabhängige Mechanismen, die Vielfalterzeugung ermöglichen.

Mit Hilfe einer Computersimulation, die auf einem Genetischen Algorithmus basiert, werden in der vorliegenden Arbeit grundsätzliche Strategien der beiden vielfaltserzeugenden Mechanismen im Rahmen einer evolutionären Entwicklung qualitativ analysiert und bewertet. Dabei wird die Genkonversion als vollständiger und die Somatische Rekombination als vereinfachter Mechanismus implementiert. Durch den Vergleich der Ergebnisse zeigt sich die Notwendigkeit von Randomisierung bei der Antikörpererzeugung für die Überlebensfähigkeit in einer sich verändernden Umwelt.

1. Abstract

The immune system of higher organisms may produce about 10^{11} [Janeway et. al., 2005, 6, p. 136] different antibodies, although the DNA with $3 \cdot 10^9$ bases encodes for less. Somatic Recombination and Gene conversion are two independent mechanisms that are responsible for the creation of diversity.

In this work the strategies of the mentioned diversity creating mechanisms will be examined and evaluated qualitatively within evolutionary development with help of computer simulation based on a Genetic Algorithm. Gene conversion is to be implemented as a whole mechanism, Somatic Recombination as a simplified one. The results indicate that randomization during the antibody generation is necessary for the survival in a changing environment.

2. Einführung

Das Immunsystem ist das Abwehrsystem höherer Organismen. Es ist ein komplexes System von Mechanismen zur Abwehr von Fremdkörpern (Antigenen).

Eine wichtige Antwort des Immunsystems auf eingedrungene Antigene ist die Sekretion von Antikörpern (Immunglobulinen). Diese binden Pathogene und deren toxische Produkte und rekrutieren weitere Zellen und Moleküle, die für die Eliminierung der Antigene sorgen. Antikörper haben eine charakteristische symmetrische Struktur:

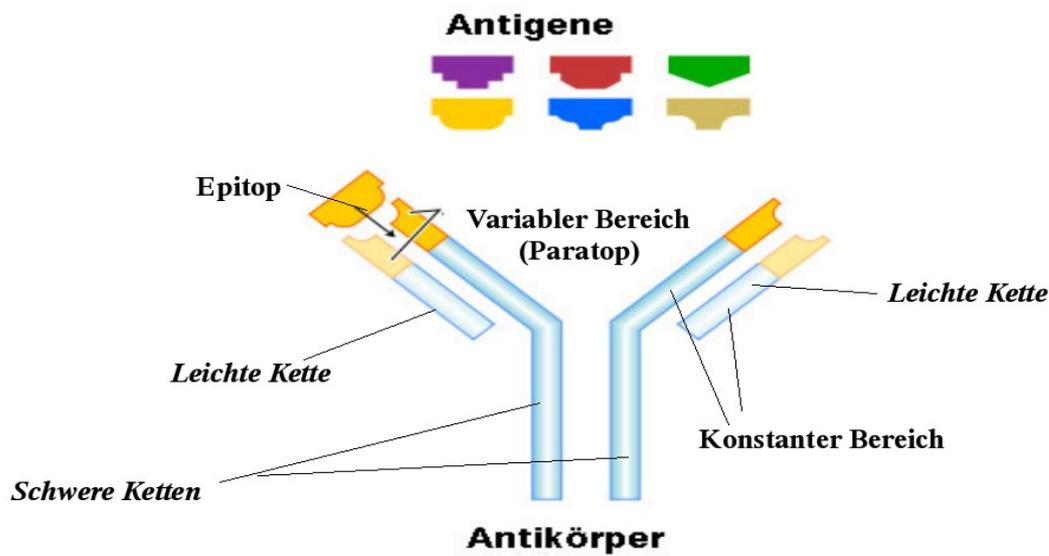


Abbildung 1: Schematische Darstellung eines Antikörpers

Antikörper bestehen aus zwei identischen schweren und leichten Ketten, welche sich in variable und konstante Domänen gliedern. Antigene werden am Epitop, einem kleinen Teil des Antigens, über die Antigenbindestelle der Antikörper, das Paratop, erkannt. Das Paratop besteht aus den variablen Domänen der schweren und leichten Kette.

Da es sehr viele unterschiedliche Antigene gibt, erzeugt das Immunsystem eine entsprechend große Zahl von Antikörpern. Diese Anzahl wird auf ca. 10^{11} [Janeway et. al., 2005, 6, S. 136] geschätzt, wobei die dafür zu Grunde liegenden Gene für sehr viel weniger Antikörper kodieren. Zur Erzeugung dieser erheblich größeren Vielfalt nutzen Spezies entweder die Somatische Rekombination (3.1) oder die Genkonversion (3.2).

Das Auftreten der beiden vielfaltserzeugenden Mechanismen stimmt nicht mit der phylogenetischen Verwandtschaft der Spezies überein. So betreiben beispielsweise Mensch und Maus Somatische Rekombination und Rinder Genkonversion, obwohl Rinder stammesgeschichtlich zwischen Maus und Mensch liegen (siehe Abbildung 2). Das langfristige Ziel ist das Verständnis der Prinzipien der Evolution des Immunsystems.

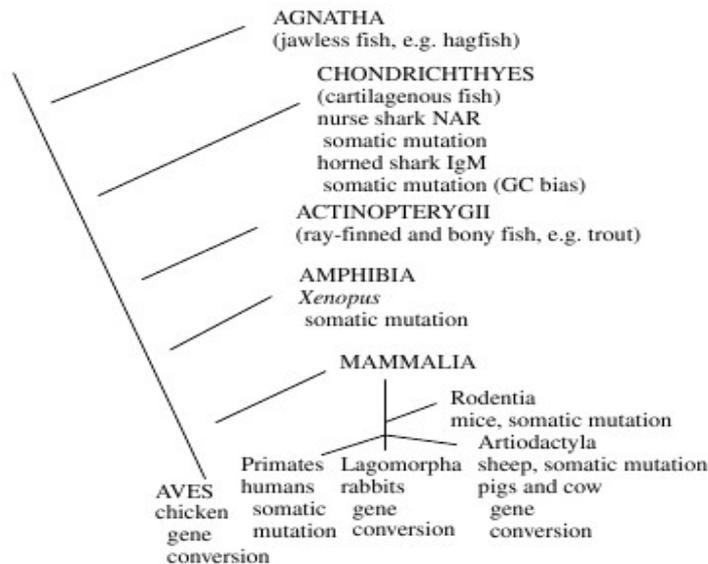


Abbildung 2: Schema der phylogenetischen Verwandtschaft von Spezies

Die phylogenetische Verwandtschaft stimmt nicht mit dem verwendeten vielfaltserzeugenden Mechanismus der Spezies überein. „Somatic mutation“ steht hier für Somatische Rekombination. [Diaz et al., 2001]

Quantitative Aussagen, beispielsweise zur Populationsgröße, sind bereits gemacht worden [Oprea and Forrest, 1998]. Allerdings ist dabei nur ein prinzipieller Mechanismus und weder die Somatische Rekombination noch die Genkonversion konkret untersucht worden. Die Arbeit von Nikolai Hecker untersucht die Genkonversion und Somatische Rekombination als einzelne Mechanismen [Hecker, 2008].

Das Ziel der vorliegenden Arbeit ist es, eine qualitative Aussage über die Eigenschaften und Unterschiede der beiden Mechanismen zu treffen. Die Untersuchungen werden mit einer Simulation durchgeführt. Dazu wird ein dem von Oprea und Forrest verwendeten Genetischen Algorithmus ähnlicher Genetischer Algorithmus implementiert. Dabei handelt es sich im Allgemeinen um ein Optimierungsverfahren: Individuen werden generiert und einer Umgebung mit Antigenen ausgesetzt, deren Fitness bestimmt und in Abhängigkeit davon eine neue Population erzeugt.

Zum Vergleich der beiden Mechanismen werden zwei Populationen erzeugt, die Vielfalt nach jeweils einem der Mechanismen generieren. Dazu wird zunächst für Somatische Rekombination (4.2.1) und Genkonversion (4.2.2) ein abstraktes Modell entworfen. Von besonderer Bedeutung ist die Erzeugung der Antikörper auf Basis der entwickelten Modelle. Auch die Proteinfaltung der Gene wird dabei berücksichtigt (4.2.3). Grundsätzlich werden sowohl die Gene und die daraus erzeugten Antikörper als auch die Antigene durch Bitstrings repräsentiert.

Die erzeugten Populationen werden einer Umgebung mit Antigenen ausgesetzt. Das Durchlaufen des Genetischen Algorithmus untersucht, ob und wie eine Anpassung der Individuen an die Antigene erfolgt und worin die Unterschiede der eingesetzten Mechanismen bestehen.

Zur Fitnessberechnung einer Population müssen die Affinitäten zwischen den Paratopen (siehe Abbildung 1) der Antikörper und den Epitopen der Antigene bzw. deren Bitstrings bestimmt werden. Dazu wird die Methode des Hammingabstandes verwendet, die den Unterschied zwischen zwei Zeichenketten ausgibt. Auf Basis der Fitnessberechnung wird im Anschluss eine neue Population aus der bestehenden erzeugt.

Durch Veränderung der Simulationsparameter wird der Einfluss jedes einzelnen untersucht (5.1) und im Anschluss daran die Simulation um zusätzliche Methoden erweitert (5.2). Auf dieser Grundlage werden Vergleiche der beiden Methoden gemacht. Auch über die Strategien der Antikörpererzeugung können Aussagen getroffen werden (6.).

Aufgrund der Tatsache, dass sich viele Untersuchungen biologisch, beispielsweise im Tiermodell, nicht realisieren lassen, sei es aus ethischen oder auch zeitlichen Gründen, bedient man sich Computersimulationen. Diese tragen zum grundsätzlichen Verständnis eines Mechanismus bei. Dabei sollte man sich immer bewusst sein, dass ein Modell der natürlichen Vorlage nie vollständig entsprechen kann. Es muss genau analysiert werden, in wie fern man auf Grund von Simulationsergebnissen auf die in der Natur vorkommenden Prinzipien schließen kann.

Da die vorliegenden Untersuchungen zu einem großen Teil auf natürlichen Vorgaben basieren, können sie einen ernstzunehmenden Beitrag für das Verständnis des Immunsystems leisten.

Gerade in Hinblick auf die Evolution sind die Untersuchungen der vielfaltserzeugenden Mechanismen mit Hilfe einer Computersimulation sinnvoll, da sich Evolution des Tierreichs in natürlicher Umgebung nicht nachstellen lässt.

3 Grundlagen der Vielfalterzeugung

Zum besseren Verständnis der entwickelten abstrakten Modelle der vielfaltserzeugenden Mechanismen im Immunsystem erfolgt in diesem Kapitel eine Erläuterung der beiden Mechanismen.

3.1 Somatische Rekombination

Die genetische Grundlage der Vielfalterzeugung bei der Somatischen Rekombination stellen die Immunglobulinsegmente des variablen Bereichs dar. Diese Genabschnitte existieren in Genclustern für zwei leichte Ketten (κ und λ) und eine schwere Kette (siehe Abbildung 1). Leichte Ketten haben je eine Region mit V- und J-Genen, schwere Ketten zusätzlich eine Region mit D-Genen. Die Anzahl und Länge der Gensegmente schwankt in Abhängigkeit von Locus und Spezies. Jeweils auf dem gleichen Chromosom sind die konstanten Genabschnitte codiert. Da sie an dem vielfaltserzeugenden Prozess keinen Anteil haben, werden sie nicht weiter betrachtet.

In den drei Clustern wird Somatische Rekombination betrieben. Dabei werden primäre RNA-Transkripte mit jeweils nur einem Segment jeder Region erzeugt. Für die leichten Ketten entstehen Transkripte mit je einem V- und ein J-Gen (VJ), für die schwere Kette ist noch ein D-Gen eingefügt (VDJ). Die Auswahl der Gene erfolgt auf Basis des Zufalls. Alle übrigen Segmente werden über einen Mechanismus herausgeschnitten (gespließt).

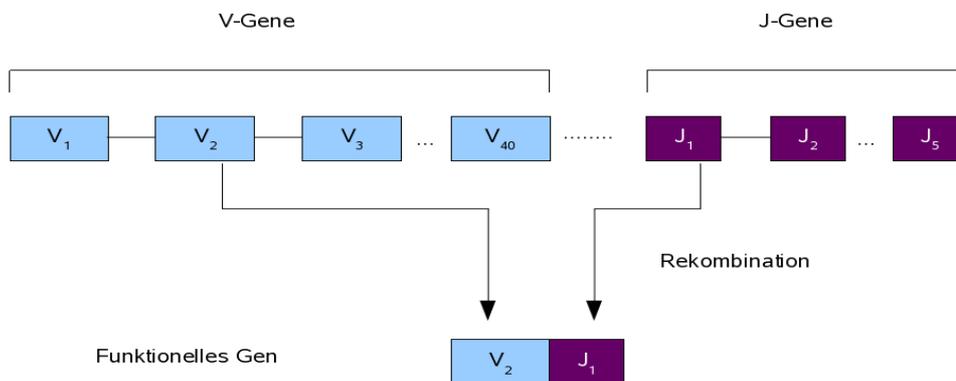


Abbildung 3: Schema der Somatischen Rekombination für eine leichte Kette κ

Somatische Rekombination erzeugt Vielfalt durch die Verknüpfung je eines zufällig ausgewählten V- und J-Genabschnitts.

Antikörpervielfalt entsteht in diesem Prozess durch eine große Anzahl von Kombinationsmöglichkeiten der Segmente der einzelnen Regionen untereinander (kombinatorische Diversität). Diese Anzahl wird durch die Verknüpfungsmöglichkeiten der unterschiedlichen leichten und schweren Ketten zusätzlich erhöht.

Damit ergibt sich die kombinatorische Vielfalt für den Menschen wie folgt:

Leichte Kette: 30 V-, 4 J- Segmente = 120 Möglichkeiten

Leichte Kette: 40 V- 5 J- Segmente = 200 Möglichkeiten

Schwere Kette: 40 V-, 25 D-, 6 J-Segmente 6000 Möglichkeiten

Daraus ergibt sich eine Vielfalt von $(120+200)*6000 = 1,9*10^6$ **verschiedenen Antikörpern**

Einen weiteren Beitrag zur Vielfalterzeugung im Rahmen der Somatischen Rekombination leistet die junktionale Diversität. Sie basiert auf der unpräzisen Verknüpfung der Segmente. Dabei können sowohl Genenden verloren gehen als auch an V- und J-Gene anschließende Genfragmente mit eingebracht werden. Ergebnis der Somatischen Rekombination ist eine große Anzahl von unterschiedlichen Primärtranskripten funktioneller Gene.

Die bis hierhin aufgeführten vielfaltserzeugenden Prozesse finden in der ersten Entwicklungsphase der B- Zellen statt. Erst auf reife B-Zellen wirkt die Somatische Hypermutation. Diese führt zahlreiche Punktmutationen in die bereits rekombinierten variablen Regionen ein, wenn ein Antigen vorhanden ist. Unter anderem Mensch und Maus betreiben Somatische Rekombination. [Janeway et. al., 2005, 6]

3.2 Genkonversion

Die Immunglobulinsegmente bei Spezies, die Genkonversion betreiben, sind ebenfalls in genetischen Loci organisiert. Die einzelnen Cluster sind allerdings anders aufgebaut: Leichte und schwere Ketten haben je nur ein V- und J-Gen, so dass durch Rekombination nur ein funktionelles Gen erzeugt werden kann. Vielfalt wird durch Pseudogene generiert. Vermutlich sind diese aus funktionellen V-Genen hervorgegangen und gleichen denen daher in ihrer Struktur. Bei Hühnern finden sich 25 Pseudogene für die leichte Kette. Während des Genkonversionsprozesses werden Fragmente mit einer Größe von 8 bis 200 Basenpaaren aus den unterschiedlichen Pseudogenen heraus kopiert. Diese Fragmente ersetzen entsprechend große Abschnitte an einer zufällig bestimmten Stelle im funktionellen Gen. Sowohl das Donorfragment als auch die Pseudogene verändern sich dabei nicht. Die Struktur des funktionellen Gens ändert sich allerdings erheblich, wobei die Länge konstant bleibt. Neben den Vögeln betreiben unter anderem auch Rinder, Schweine und Hasen Genkonversion. [Arakawa and Buerstedde, 2004]

Dadurch dass die Genkonversion auf einem komplexen Zusammenspiel von zufälligen Mechanismen basiert, ist die Anzahl an möglichen Primärtranskripten schwer abzuschätzen. Wie bei der Somatischen Rekombination erhöht sich die Anzahl unterschiedlicher Primärtranskripte zusätzlich noch durch die Kombinationsmöglichkeiten der unterschiedlich generierten leichten und schweren Ketten.

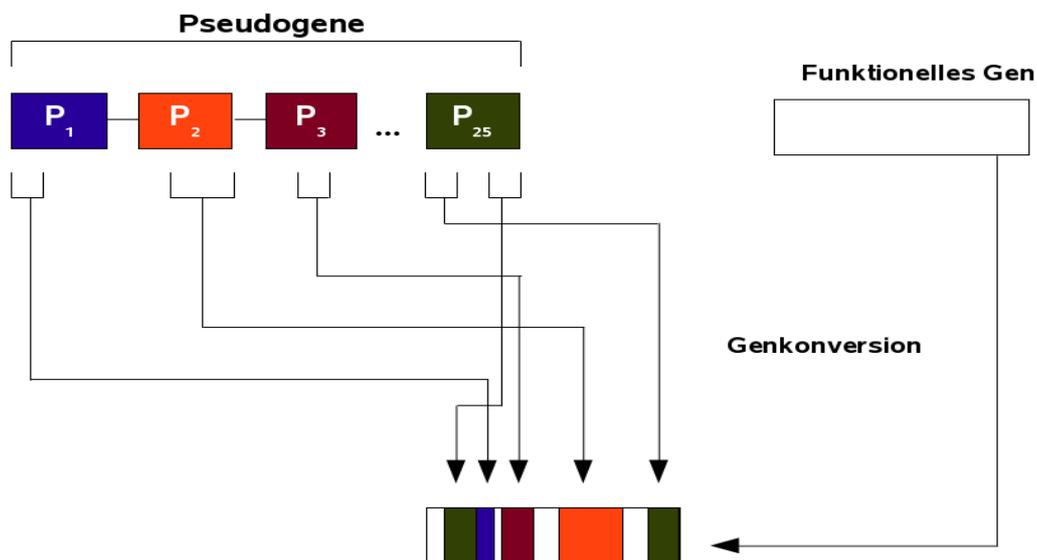


Abbildung 4: Schema der Genkonversion für eine leichte Kette

Genkonversion beruht auf drei zufällig bestimmten Komponenten: Die Größe und die Position der Pseudogenabschnitte, die kopiert werden, und der Abschnitt, der in dem funktionellen Gen ersetzt wird.

4. Simulation

In diesem Kapitel wird auf die Struktur und konkrete Umsetzung des verwendeten Genetischen Algorithmus so wie die nach dem Vorbild der in Kapitel 3 erläuterten grundsätzlichen Prinzipien entwickelten Modelle der vielfaltserzeugenden Mechanismen eingegangen.

4.1 Genetischer Algorithmus

Um die Somatische Rekombination und Genkonversion in Bezug auf die Evolution zu untersuchen, wird ein genetischer Algorithmus verwendet.

Das grundlegende Prinzip ist die Erzeugung einer Ausgangspopulation und das anschließende Durchlaufen einer festgelegten Anzahl an Generationen. Zu Beginn jeder Generation bzw. jedes Durchlaufes wird die Fitness bestimmt. Auf Basis dieser werden Individuen für die neue Generation erzeugt.

Der verwendete Algorithmus basiert in der allgemeinen Struktur auf dem von Oprea und Forrest in „Simulated evolution of antibody gene libraries under pathogen selection“ (1998) publizierten Algorithmus (siehe Abbildung 5). Die detaillierte Struktur und Umsetzung des Algorithmus wird in den folgenden Abschnitten 4.2 bis 4.4 erläutert.

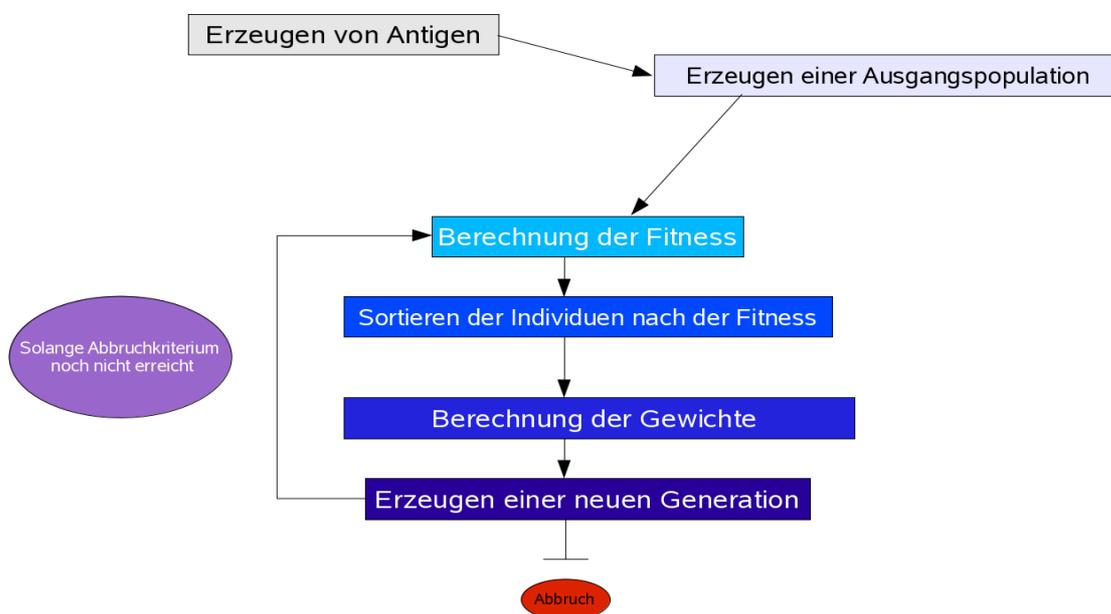


Abbildung 5: Schema des Genetischen Algorithmus nach Oprea und Forrest [Oprea and Forrest, 1998]

Ein Genetischer Algorithmus ist ein Optimierungsverfahren, dessen Grundstruktur meist sehr ähnlich ist: Individuen werden generiert, deren Fitness bestimmt und in Abhängigkeit davon eine neue Population erzeugt.

4.2 Modelle der vielfaltserzeugenden Mechanismen

Die Individuen aller Populationen (Ausgangs- und Folgepopulationen) und die von ihnen generierten Antikörper werden anhand eines Modells der beiden vielfaltserzeugenden Mechanismen erzeugt. Das entwickelte Modell der Somatischen Rekombination und der Genkonversion wird in den folgenden Abschnitten erläutert.

4.2.1 Modell der Somatischen Rekombination

Die Somatische Rekombination besteht aus den drei vielfalterzeugenden Prozessen kombinatorische und junktionale Diversität, sowie der Somatischen Hypermutation. In dem Modell wird nur die kombinatorische Diversität berücksichtigt. Durch die Untersuchungen soll geklärt werden, aus welchen Gründen zwei weitere Mechanismen zur Vielfalterzeugung notwendig sind. Das Modell ist vereinfacht, da nur für die leichte Kette κ der Prozess betrachtet wird.

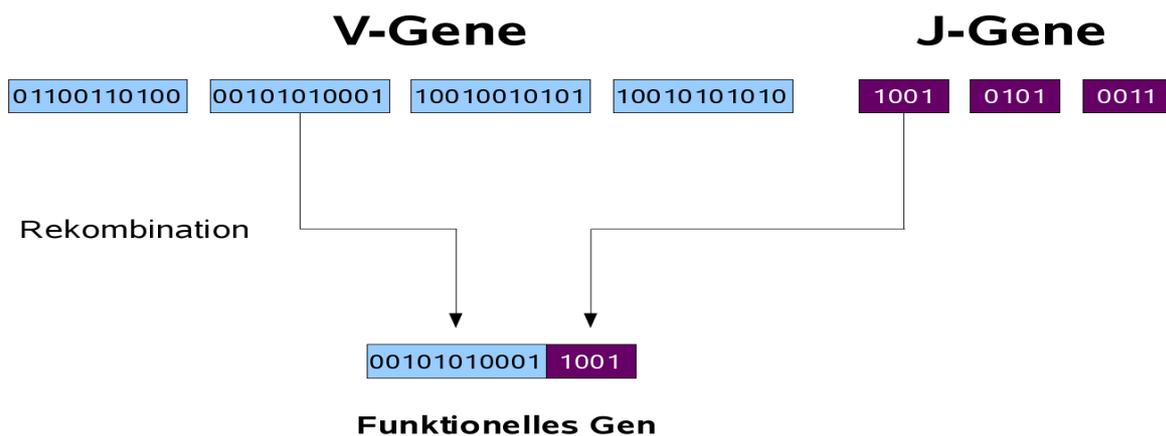


Abbildung 6: Modell der Somatischen Rekombination

Somatische Rekombination erzeugt Vielfalt durch die Verknüpfung je eines zufällig ausgewählten V- und J-Genabschnitts. Die dargestellten Längenverhältnisse entsprechen nicht den natürlichen (siehe 4.2.1).

Gemäß der natürlichen Vorlage haben die in der Simulation erzeugten Individuen V- und J-Gene für die leichten Ketten. Die Längen dieser Gene werden der ImMunoGeneTics-Datenbank entnommen [<http://www.imgt.org/>]. Da unter anderem die Maus Somatische Rekombination betreibt, verwendet man die Durchschnittswerte für die Längen der V- und J-Gene der leichten Kette κ bei der Maus für das Modell. Damit ergeben sich 40 V-Gene mit einer Länge von 300 Basenpaaren und 5 J-Gene mit einer Länge von 40 Basenpaaren.

Alle Gene sowie die daraus gebildeten Antikörper werden durch zufällig erzeugte Bitstrings repräsentiert. V- und J-Gen-Cluster werden über eine Methode erzeugt, die einen Array der Größe m mit Bitstrings der Länge n zurück gibt. Zur Generierung funktioneller Gene werden per Zufall je ein V- und ein J-Gen ausgewählt und verknüpft. Ergebnis dieses Mechanismus ist eine Bibliothek von Genen einer Länge von 340 Bits.

Hier ist zu beachten, dass eigentlich 2 Bits zur Repräsentation der vier Basen der DNA notwendig sind. Da der Informationsgehalt der V- und J-Segmente niedriger ist als die tatsächliche Länge, ist es ausreichend, wenn in der Simulation die Relationen der Genlängen berücksichtigt werden.

4.2.2 Modell der Genkonversion

Genkonversion erzeugt Vielfalt durch Modifikation. Dabei nehmen die zufällig bestimmten Werte der Größe und Position des ausgewählten Pseudogenabschnitts und der in dem funktionellen Gen ersetzte Abschnitt Einfluss auf den Umfang der Veränderung. Auch hier ist das Modell vereinfacht, da der Prozess ebenfalls nur für eine leichte Kette betrachtet wird.

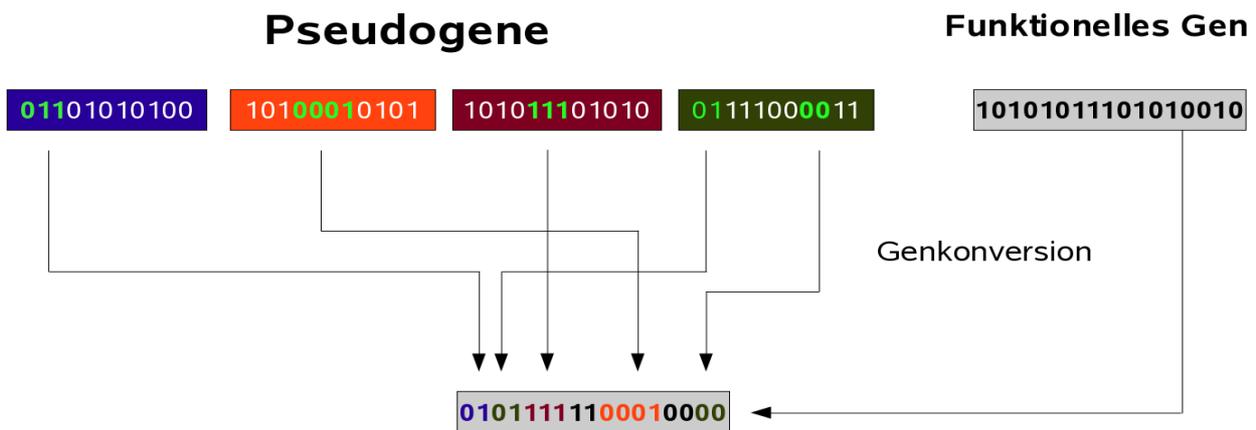


Abbildung 7: Modell der Genkonversion

Genkonversion beruht auf drei zufällig bestimmten Komponenten: Die Größe und die Position der Pseudogenabschnitte, die kopiert werden, und der Abschnitt, der in dem funktionellen Gen ersetzt wird. Die dargestellten Längenverhältnisse entsprechen nicht den natürlichen (siehe 4.2.2).

Im Gegensatz zu der natürlichen Vorgabe haben die durch das Programm erzeugten Individuen ein funktionelles Gen anstelle je eines V- und J-Gens, da durch Verknüpfung von V- und J-Gen nur ein funktionelles Gen entstehen kann. Es werden entsprechend der natürlichen Vorgabe bei Hühnern 25 Pseudogene erzeugt. Sowohl das funktionelle Gen als auch die Pseudogene werden durch eine Methode generiert, die zufällige Bitstrings der Länge m erzeugt.

Da das durch die Somatische Rekombination erhaltene funktionelle Gen eine Länge von 340 Basenpaaren umfasst, wird diese Länge ebenfalls für das funktionelle Gen bei der Genkonversion gewählt. Für die Genkonversion bei Hühnern ist bekannt, dass maximal 200 Basenpaare aus den Pseudogenen kopiert werden können [Arakawa and Buerstedde, 2004]. Daher werden Pseudogene mit einer Länge von 200 Bits für die Simulation erzeugt.

Der Literatur sind bezüglich der Genkonversion keine exakten Werte für die Auswahl der Pseudogenfragmente, die ersetzten Positionen im funktionellen Gen und die Anzahl der Genkonversionsevents zu entnehmen. Damit die Wahrscheinlichkeit, in das funktionelle Gen eingeführt zu werden, für jede Base der Pseudogene gleich ist, werden die benötigten Parameter zufällig bestimmt. Auch für die Anzahl der Genkonversionsevents wird eine Zufallszahl bestimmt. Ein Parameter legt dafür die maximal mögliche Anzahl der Events fest. Ergebnis dieses Mechanismus ist ebenfalls eine Genbibliothek.

4.2.3 Modell der Antikörpergenerierung

In der Natur entsprechen die Gene nicht den daraus gebildeten Proteinen, hier also Antikörpern. Sie sind beliebig kompliziert gefaltet, wodurch nur ein Teil der Basensequenz als Kontaktfläche zwischen Antigen und Antikörper dient. Auch die Gene in der Simulation werden einem „Faltungsmechanismus“ unterzogen. Der Mechanismus kopiert Teilbereiche aus einem zuvor erzeugten Gen, wodurch Antikörper unterschiedlicher Länge entstehen. Auf Grund der zufälligen Auswahl der Teilbereiche ist gewährleistet, dass sich alle Gene in Abhängigkeit ihrer Sequenz unterschiedlich falten und bei identischer Sequenz immer gleich. Dies entspricht der natürlichen Vorstellung. In der Natur kann es vorkommen, dass sich Proteine nur um eine Base unterscheiden, aber eine völlig andere räumliche Struktur einnehmen. Dies gewährleistet der Mechanismus nicht.

Pseudocode für Antikörpergenerierung:

input

gene : gene, randomly generated bitstring (class Individuum)

new_paratop : " "

counter: 0

output

new_paratop: created antibody

for each *gene* [i] **do**

if (*gene* [i] = 1) **then**

counter \leftarrow *counter* +1

if (*counter* = 2 or *counter* = 3) **then**

new_paratop \leftarrow *new_paratop* + *gene* [i]

if (*counter* = 14) **then**

counter \leftarrow 0

Der *counter* ist auf Basis der Überlegungen zur Affinitätsformel und der Antikörperlänge festgelegt worden (siehe 4.2.1). Ergebnis dieser Faltungsmethode sind Antikörper einer durchschnittlichen Länge von 23 Bits.

4.2.4 Repräsentation der Antigene

Da die Antikörper im weiteren Verlauf der Simulation mit den Antigenen verglichen werden, generiert die Methode, die einen Array der Größe *m* mit Bitstrings der Länge *n* erzeugt, eine Antigenbibliothek. Die Länge und die Anzahl sind prinzipiell frei wählbar, da es keine Informationen in der Literatur gibt. In den vorliegenden Simulationen ist diese auf die Länge und Anzahl von Antikörpern abgestimmt worden.

4.3 Fitnessberechnung

Nachdem eine Population an Individuen und eine Antigenbibliothek gemäß der Modelle erstellt worden sind, setzt sich der Durchlauf des Genetischen Algorithmus mit der Fitnessbestimmung fort. Dazu werden zunächst die Affinitäten zwischen Antikörpern und Antigenen bestimmt. Dies geschieht für jeden Antikörper aller Individuen einer Population.

Zur Berechnung der Affinität von zwei Bitstrings wird die Methode des maximalen Hammingabstandes verwendet. Dabei erhält man den größten Unterschied der beiden Sequenzen dadurch, dass der kürzere String über den längeren geschoben und jeweils die Anzahl der ungleichen Bits bestimmt wird.

Folgende Formel repräsentiert die Bestimmung des maximalen Hammingabstandes D:

$$D = \max_j \sum_{i=1}^L \delta(\text{Antikörper}_i, \text{Antigen}_{i+j}) \quad \text{mit } L = \text{Länge des Antikörpers} \quad (1)$$

Um einen Affinitätswert zwischen '0' und '1' zu erhalten, wird der maximale Abstand durch die Länge des kürzeren Strings (Antikörpers) geteilt.

Aus der Formel erhält man eine lineare Gleichverteilung für die Werte. Dies entspricht nicht der natürlichen Vorstellung, da die Affinität zwischen einem Antigen und einem Antikörper nicht basenweise zunimmt. Bis zu einer bestimmten Anzahl an gleichen Basenpaaren findet überhaupt keine Erkennung statt. Ab einer Schwelle kann jede weitere komplementäre Base jedoch einen zunehmend größeren Beitrag zu einer höheren Affinität leisten. Durch das Einführen eines Exponenten wird dies in die Simulation eingebracht. Der Exponent wird nach folgenden Überlegungen festgelegt:

Man geht davon aus, dass ein Antikörper mit einer durchschnittlichen Affinität von 10^{-6} an ein Antigen bindet [Blattman et al., 2002]. Für die Erkennung jedes Antigens benötigt dafür man rechnerisch eine Antikörperanzahl von 10^6 . Für die Generierung dieser Antikörperanzahl benötigt man mindestens 20 Bits. Bei der Berechnung der Binomialverteilung auf Basis unterschiedlicher Antikörperlängen und verschiedener Werte für den Exponenten stellt man eine Abhängigkeit beider Werte fest: Je länger der generierte Antikörper ist, desto niedriger kann der Exponent sein, um dennoch einen Durchschnittswert von 10^{-6} für die Affinität zu erhalten. Prinzipiell ist es unerheblich, ob man einen größeren Exponenten und eine geringe Bitanzahl wählt oder umgekehrt. In Hinblick auf das mögliche Einführen von Mutationen zu einem späteren Zeitpunkt, ist die Generierung eines längeren Antikörpers sinnvoll. So können Mutationen einen weniger großen Einfluss haben. Bei der

Simulation wird der Exponent auf 40 gesetzt:

$$\text{Affinität} = \left(\frac{D}{L}\right)^{40} \quad \text{mit } L = \text{Länge des Antikörpers} \quad (2)$$

Der in der Methode zur Antikörpergenerierung (siehe 4.1.3) eingeführte Counter wird auf 14 gesetzt, damit die erzeugten Antikörper ungefähr eine durchschnittliche Länge von 20 Bits haben.

Mit Formel 2 werden für alle Antikörper eines jeden Individuums die Affinitäten gegenüber der Antigene bestimmt. Diese geben für jedes Individuum an, wie gut es an das entsprechende Antigen angepasst ist. Die Fitnesswerte erhält man aus den Affinitätswerten. Dazu werden die Gegenwahrscheinlichkeiten (1- Affinitätswert) der Affinitäten miteinander multipliziert und von der maximalen Fitness ('1') subtrahiert.

Als Fitnesswert für ein Individuum wird jeweils die geringste Fitness angenommen. Dies wird angenommen, da schlechte Anpassung an ein Antigen zum „Sterben“ ausreicht.

4.4 Generierung einer neuen Population

Bevor eine neue Generation erzeugt werden kann, müssen nach Vorgabe des Genetischen Algorithmus einige Schritte für die aktuelle Population durchlaufen werden. Zur Bestimmung ihrer relativen Wahrscheinlichkeiten für die Fortpflanzung werden die Individuen fitnessabhängig sortiert. Zur Bestimmung der Wahrscheinlichkeiten oder auch Gewichte werden die Individuenanzahl M und der Rang r eines Individuums in folgende Gewichtsformel eingesetzt:

$$W = 2 \cdot (M - r) / (M \cdot (M - 1)) \quad (3)$$

Als Ergebnis der Anwendung der Gewichtsformel erhält jedes Individuum einen Gewichtswert zwischen '0' und '1', wobei sich diese zu '1' aufaddieren. Aus den Gewichtswerten ergibt sich eine Wahrscheinlichkeitsverteilung, wodurch ein Individuum mit einer hohen Fitness einen größeren Wahrscheinlichkeitsbereich als Individuen mit einer niedrigeren Fitness erhält.

M Individuen für eine neue Generation werden nach folgendem Schema erzeugt:

Mit einem Zufallsgenerator wird für beide Elternindividuen eine Zahl zwischen '0' und '1' erzeugt. Anhand der zuvor aufgestellten Wahrscheinlichkeitsverteilung wird jeweils das entsprechende Individuum ausgewählt. Per Zufall wird festgelegt, welches Individuum der Vater und welches die Mutter ist. Bis zu einem zufällig bestimmten Punkt, dem Crossingover-Punkt werden die Gene vom Vater übernommen, danach von der Mutter. Der Crossingover-Punkt liegt immer zwischen V- und J-Genen bzw. Pseudogenen und funktionellem Gen, so dass nur ganze Gene übertragen werden. Die neu erzeugte Generation wird mit einer Mutationsrate von 0,02% mutiert. Bis zu einer definierten Anzahl von Generationen startet der Algorithmus wiederum mit der Fitnessberechnung. Bei den Simulationen im Rahmen dieser Arbeit ist die Anzahl der Generationsdurchläufe auf 100 festgelegt worden, weil die Fitness in allen durchgeführten Simulationen gegen ihren maximalen Wert konvergiert ist.

4.5 Implementierung

Die Implementierung der Simulation ist in Java erfolgt. Dafür sind unterschiedliche Klassen erstellt worden: Eine Klasse enthält Methoden zur Generierung von zufälligen Zahlen und Bitstrings, Methoden zur Affinitätsbestimmung und eine Methode zur Einführung von Punktmutationen. Die Klassen der Individuen, Antikörper und Antigene enthalten Methoden entsprechend ihrer „Fähigkeiten“, z.B. können Individuen Antikörper erzeugen oder Antikörper gefaltet werden. In einer Hauptklasse ist die Simulation des Genetischen Algorithmus zu finden. Von dort aus werden entsprechend des erläuterten Ablaufs des Algorithmus (4.1-4.4) die anderen Methoden aufgerufen.

5. Ergebnisse

In diesem Kapitel werden die Ergebnisse der Simulationen objektiv analysiert. Die Simulationen durchläufe stellen die Grundlage der Diskussion (6.) und Bewertung der Methoden dar. Man verschafft sich zunächst einen Überblick durch das Einsetzen der in Tabelle 1 (siehe Anhang) aufgeführten Standardwerte und verändert im Anschluss daran die nicht durch die Literatur vorgegebenen Parameter. Nach der Implementierung zusätzlicher Methoden schließen sich weitere Untersuchungen an.

5.1 Versuche auf Basis von Standardbedingungen

5.1.1 Fitnessentwicklung bei Genkonversion und Somatischer Rekombination

Erste Erkenntnisse erhält man aus einer Simulation beider Methoden unter Standardbedingungen:

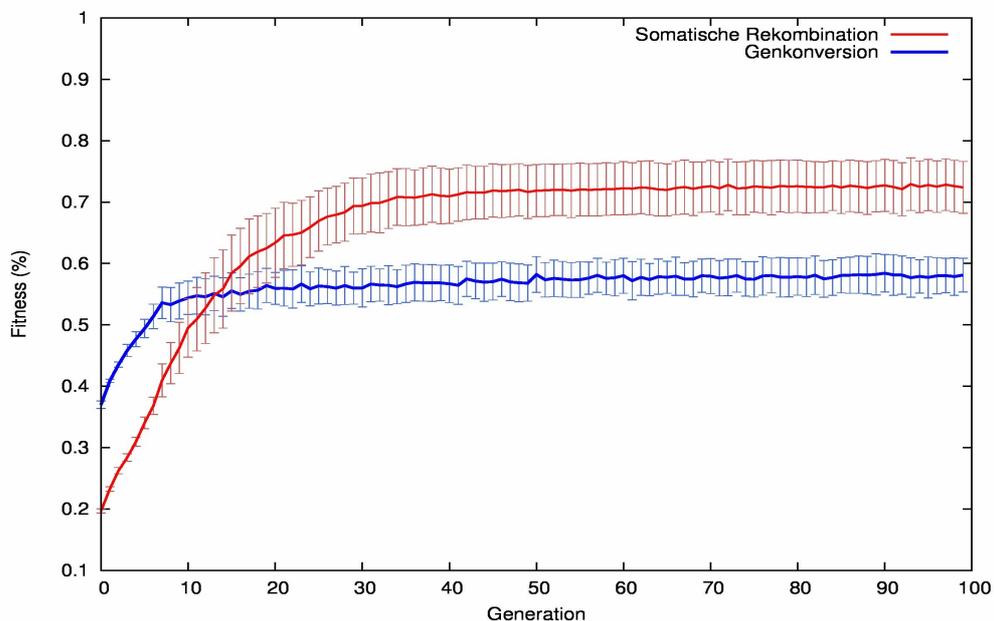


Abbildung 8: Die beiden Methoden im einfachen Vergleich; Parameter: Standardwerte gemäß Tabelle 1; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Diese zeigt Unterschiede der beiden Mechanismen auf. Auffällig ist die Ausgangsfitness der beiden Populationen: Individuen, die Genkonversion betreiben, erreichen eine höhere Startfitness als Individuen, die Somatische Rekombination betreiben. Charakteristisch sind ebenfalls die Entwicklungen der Fitnesswerte: Sie steigen bei gekonversionbetreibenden Individuen kontinuierlich und erreichen bereits nach ca. 25 Generationen ihren maximalen Wert von ca. 0,57. Es findet eine Fitnesssteigerung um ca. 20% statt. Schwankungen der Fitnesswerte sind auffällig. Im Gegensatz dazu steigen die Fitnesswerte bei Individuen, die Somatische Rekombination betreiben, über einen Zeitraum von 35 Generationen um ca. 50%. Die Schwankungen um den erreichten maximalen Fitnesswert von ca. 0,7 sind wenig ausgeprägt.

5.1.2 Einfluss der Populationsgröße

Da einige der angenommenen Standardparameter nicht auf Literaturwerten basieren, muss deren Einfluss auf die Mechanismen untersucht werden. Nur so kann man falsche Schlüsse auf Basis der angenommenen Werte ausschließen. Dies dient zudem der Präzision der Methodeigenschaften.

Im ersten Versuch ist die Simulation mit einer genügend großen Population durchgeführt worden. Aus Untersuchungen zur Evolution [Campbell and Reece, 2003, 6, S. 503-569] ist bekannt, dass Weiterentwicklung im Rahmen unterschiedlich großer Populationen statt gefunden hat. Verringerungen der Populationsgröße sind unter anderem durch räumliche Trennung oder Umwelteinflüsse ausgelöst worden. Interessant zu untersuchen ist daher, welchen Einfluss die Populationsgröße auf die Mechanismen hat:

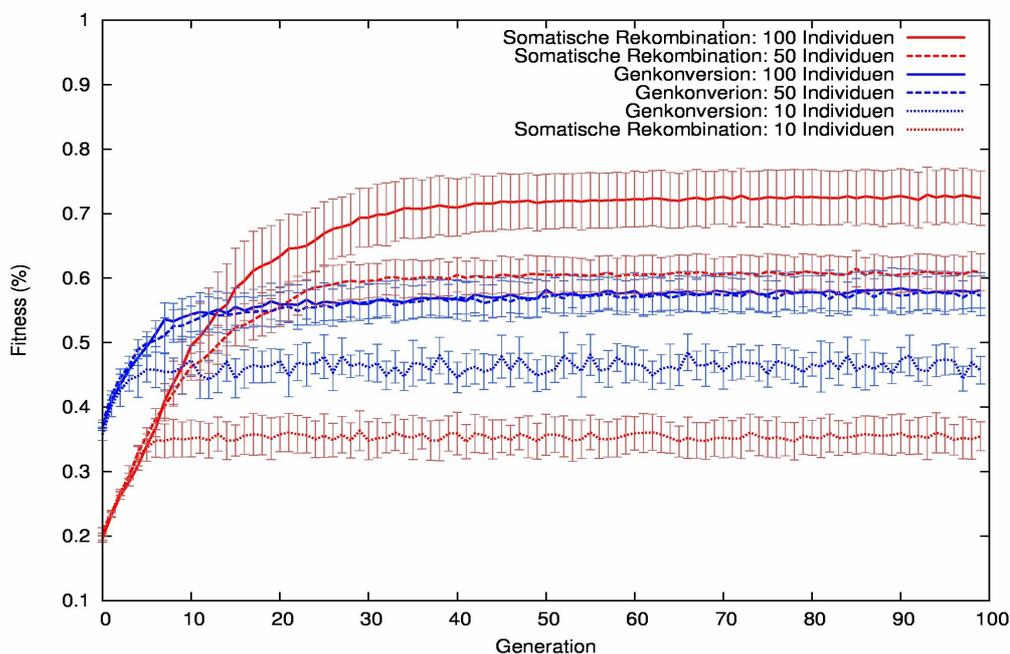


Abbildung 9: Einfluss der Populationsgröße auf die Fitnessentwicklung; Parameter: Anzahl von Individuen: 100; 50; 10; Übrige Werte gemäß Tabelle 1; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Während sich die maximale Fitness bei genkonversionbetreibenden Individuen bei einer Verringerung der Individuenanzahl um 50% nicht signifikant ändert, sinkt diese bei rekombination betreibenden Individuen um mehr als 10%. Bei einer weiteren Reduzierung der Populationsgröße auf 10% der Ursprungsgröße weisen auch die Individuen, die Genkonversion betreiben, einen um ca. 10% geringeren und signifikant verschiedenen Fitnesswert auf. Unter diesen Bedingungen erzielen sie eine höhere maximale Fitness als rekombinationbetreibende Individuen. Ihr maximaler Fitnesswert ist insgesamt um mehr als 35% gesunken. Auffällig ist die Unveränderlichkeit der Ausgangsfitnesswerte.

5.1.3 Einfluss der Antikörperanzahl

Mit der Verringerung der Populationsgröße gehen die Senkung der Gesamtgenanzahl und damit die Reduzierung der theoretischen Vielfalt innerhalb einer Population einher. Daher wird in der folgenden Simulation der Einfluss der Antikörperanzahl durch die Verringerung der Antikörperanzahl untersucht:

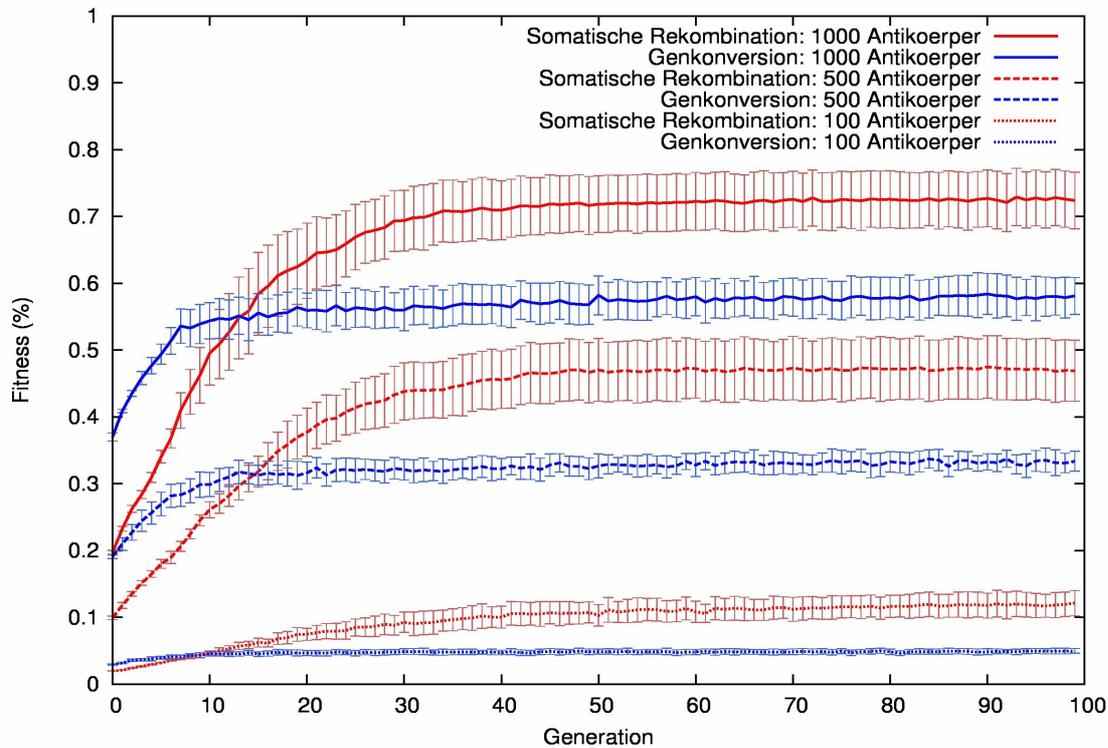


Abbildung 10: Einfluss der Antikörperanzahl auf die Fitness; Parameter: Anzahl an Antikörpern: 1000, 500, 100; Übrige Werte gemäß Tabelle 1; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Das Ergebnis der Simulation zeigt, dass eine Verringerung der Antikörperanzahl auf die Somatische Rekombination und Genkonversion Einfluss hat. Bei beiden Mechanismen sinken Startfitness und maximale Fitness. Die Startfitness der Somatischen Rekombination kann sich bei einer Verdoppelung der Antikörperanzahl im Gegensatz zu der Genkonversion nur um ca. 10% steigern. Bei dieser findet eine Erhöhung der Anfangsfitness um ca. 25% statt. Das relative Verhältnis der Fitnesswerte bleibt allerdings erhalten. Die Somatische Rekombination erreicht nach wie vor höhere Fitnesswerte als die Genkonversion. Auch der Kurvenverlauf bleibt unverändert.

5.1.4 Einfluss der Antigenanzahl

Gegenspieler der Antikörper im Immunsystem sind die Antigene. Je besser die Antikörper die Antigene erkennen, desto höher ist die Fitness eines Individuums. In den vorangegangenen Versuchen wurde gezeigt, dass eine zunehmende Antikörperanzahl einen positiven Effekt auf die Fitness hat. Dabei ist die Anzahl der Antigene unverändert geblieben. Anzunehmen ist daher, dass die Erhöhung der Antigene ebenfalls einen Einfluss auf die Fitness hat. Dies wird in einer weiteren Simulation getestet.

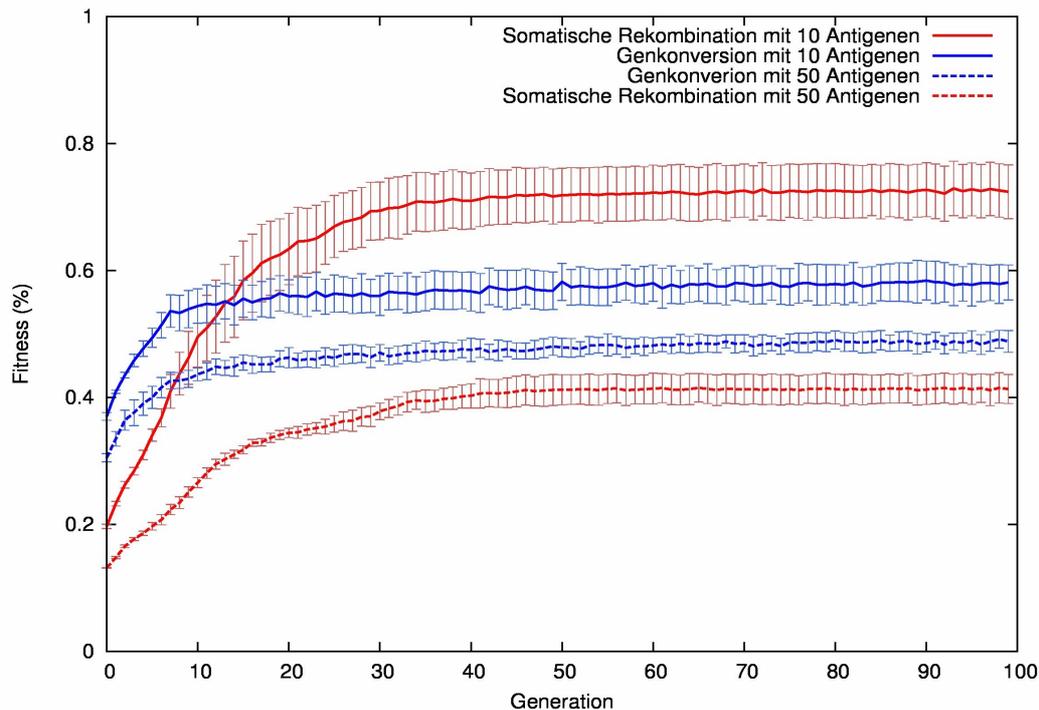


Abbildung 11: Einfluss der Antigenanzahl auf die Fitness; Parameter: Anzahl an Antigenen: 10 und 50; Übrige Werte gemäß Tabelle 1; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Die Antigenanzahl hat Einfluss auf die Fitness. Die Fitness der Startpopulation und die maximale Fitness sinken bei erhöhtem Antigenvorkommen. Die Genkonversion wird weniger stark beeinflusst als die Somatische Rekombination. Individuen, die Genkonversion betreiben, büßen nur ca. 10% ihrer Fitness ein, recombination betreibende Individuen hingegen ca. 30%. Damit weisen gekonversion betreibende Individuen bei einer größeren Antigenanzahl eine höhere Fitness auf als Individuen, die Somatische Rekombination betreiben.

5.1.5 Einfluss der Antigenlänge

Die Größe der Antikörper ist relativ gut durch die Literaturwerte gesichert (siehe Tabelle 1, Anhang). Die tatsächliche Größe eines Antigens ist allerdings nicht bestimmbar. Da Antigene eine sehr große inhomogene Gruppe bilden, weisen sie sehr unterschiedliche Strukturen und Größen auf. Es ist bekannt, dass Antikörper Antigene oft nur in einem kleinen Bereich an der Oberfläche erkennen. Die Veränderung der Antigenlänge soll Aufschluss über ihren Einfluss auf die Fitness geben.

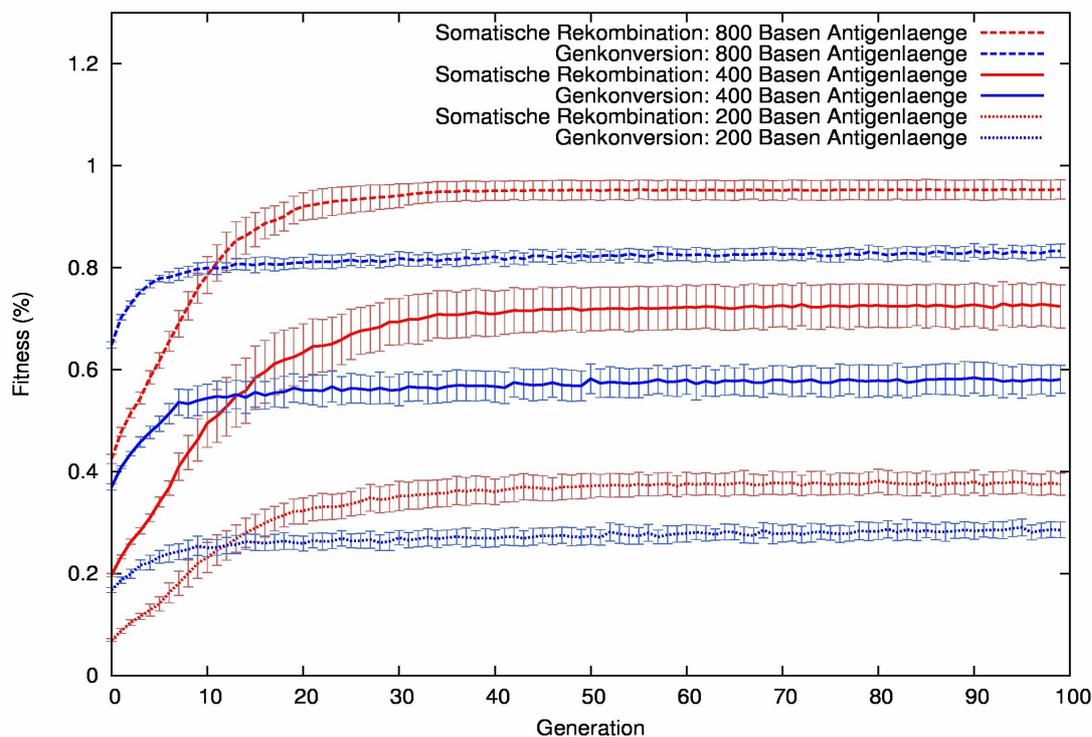


Abbildung 12: Einfluss der Antigenlänge auf die Fitness; Parameter: Länge der Antigene: 200, 400 und 800 Bits; Übrige Werte gemäß Tabelle 1; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Die Simulationen zeigen, dass die Antigenlänge Einfluss auf die Fitness hat: Unabhängig vom Mechanismus führt eine größere Antigenlänge sowohl zu einer höheren Startfitness als auch maximalen Fitness und umgekehrt. Die Somatische Rekombination bedingt weiterhin eine insgesamt höhere Fitness.

5.1.6 Einfluss der Anzahl an Genkonversionsevents

In der Natur wird die Somatische Rekombination um die junktionale Diversität und Somatische Hypermuation bereichert. Die Genkonversion ist bei der Vielfalterzeugung allein auf die Genkonversionsevents angewiesen. Da die Anzahl an Genkonversionsevents unbekannt ist, wird deren Bedeutung durch Veränderung der Eventanzahl in einer weiteren Simulation untersucht.

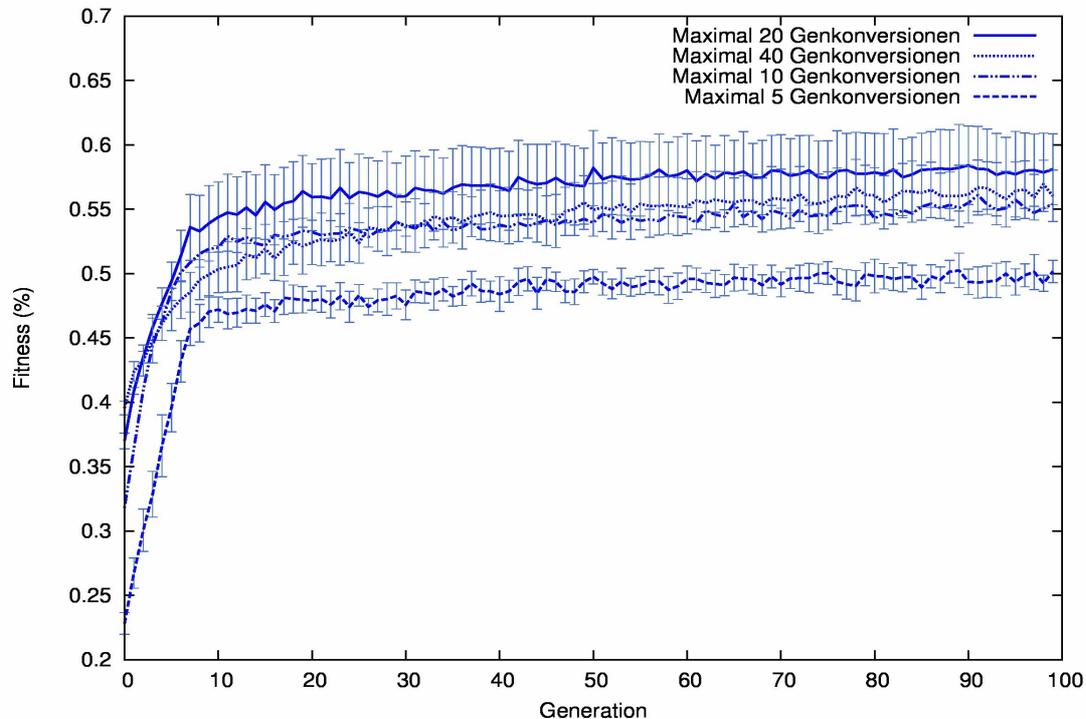


Abbildung 13: Einfluss der Anzahl an maximal möglichen Genkonversionsevents auf die Fitness

Parameter: Maximale Anzahl an Genkonversionsevents: 5, 10, 20, 40; Übrige Werte gemäß Tabelle 1; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Die Ergebnisse zeigen, dass die Anzahl der Genkonversionsevents Einfluss auf die Fitness nehmen. Sowohl die Ausgangsfitness als auch die maximale Fitness ändern sich in Abhängigkeit der Anzahl an Genkonversionen. Je höher die Anzahl an Genkonversionen, desto höher ist die Ausgangsfitness. Für die maximale Fitness trifft das nicht zu. Individuen, die nur 5 Genkonversionen durchführen können, erreichen eine geringere maximale Fitness als Individuen, die 10 bis 40 Genkonversionen durchführen. Zwischen diesen zeigt sich kein signifikanter Unterschied. Es kann eine größere Fitnesssteigerung beobachtet werden, wenn die Anzahl an Genkonversionsevents geringer ist.

5.1.7 Einfluss der Generierung einer neuen Antigenbibliothek

Bis zu diesem Zeitpunkt sind alle Parameter während einer Simulation unverändert geblieben und Entwicklung hat ausschließlich durch natürliche Selektion statt gefunden. Dabei stehen Fitness und Reproduktion in einem direkten Verhältnis. Evolution hat aber auch durch schlagartige Änderung der Umweltbedingungen statt gefunden. Dabei werden die Populationen plötzlich vollständig neuen Bedingungen ausgesetzt [Campbell and Reece, 2003, 6, S. 503-569].

In der folgenden Simulation werden neue Umweltbedingungen dadurch erzeugt, dass nach unterschiedlich vielen Generationen komplett neue Antigene erzeugt werden.

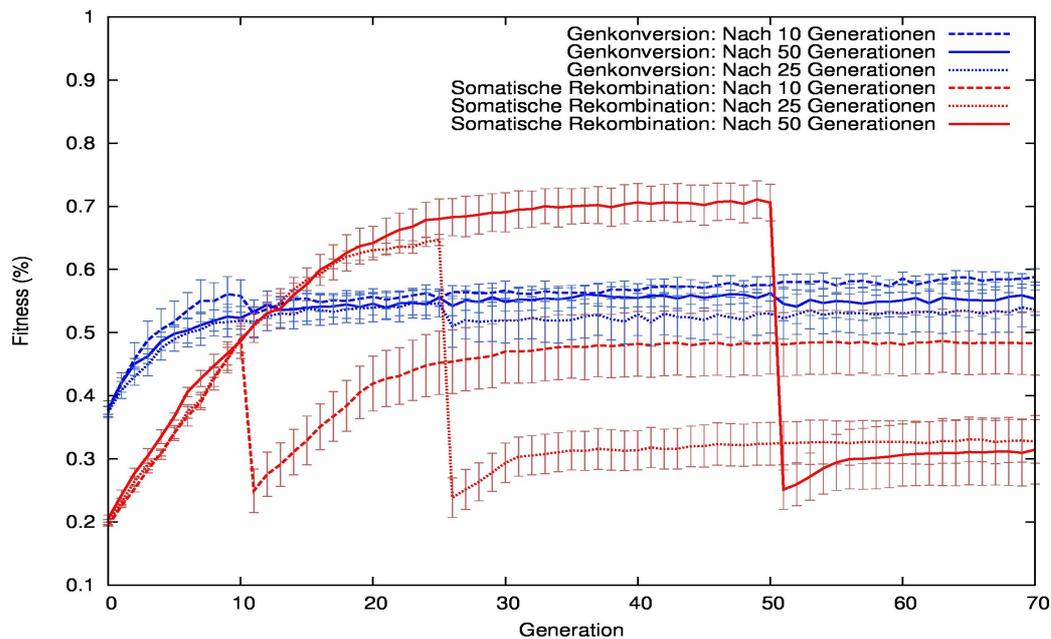


Abbildung 14: Einfluss der Generierung neuer Antigene nach unterschiedlich vielen Generationen auf die Fitness; Parameter: Werte gemäß Tabelle 1; Nach 10, 25 und 50 Generationen werden neue Antigene erzeugt; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen; Es wurden ebenfalls 100 Generationen durchlaufen, wovon nur 70 gezeigt werden

An den Simulationen kann zweierlei beobachtet werden: Zum einen werden die beiden Mechanismen ganz unterschiedlich beeinflusst, zum anderen hat der Zeitpunkt der Generation neuer Antigene vor allem Einfluss auf die Somatische Rekombination. Die Neugeneration von Antigenen beeinflusst die Genkonversion gering. Kurz nach der Generierung ändert sich die Fitness um weit weniger als 10%. Bei der Somatischen Rekombination verringert sich die maximale Fitness um bis zu 45% nach der Neugeneration. Auch der Zeitpunkt der Generation der neuen Antigene ist von entscheidender Bedeutung. Zwar fällt die maximale Fitness auf immer ungefähr denselben Wert von ca. 0,25 zurück, dennoch kann sich die Fitness auf einen höheren Wert steigern, wenn die Neugeneration zu einem früheren Zeitpunkt statt gefunden hat.

5.2. Versuche auf Basis von Simulationserweiterungen

Bislang sind alle Versuche auf Basis der entwickelten Simulation durchgeführt worden. Es lassen sich bereits Aussagen über die Methoden treffen. Zur Konkretisierung der Bewertung erfolgen zusätzliche Veränderungen.

5.2.1 Einfluss der kontinuierlichen Erzeugung neuer Antigene

Da die Neugeneration von Antigenen beide Mechanismen in unterschiedlicher Weise beeinflusst und gerade die Veränderungen der äußeren Bedingungen eine wichtige Aufgabe im Rahmen der Evolution einnehmen, setzt die Simulationsveränderung dort an. In 5.1.7 ist die Neugeneration einer kompletten Antigenbibliothek untersucht worden. Realistischer ist eine geringere Veränderung der äußeren Umgebung. Daher wird die Simulation um eine Methode erweitert, die den Austausch eines Teils der Antigene sowie eine kontinuierliche Veränderung der Antigene ermöglicht. Zunächst sind zur weiteren Untersuchung nach jeweils 10 Generationen 10% neue Antigene erzeugt worden:

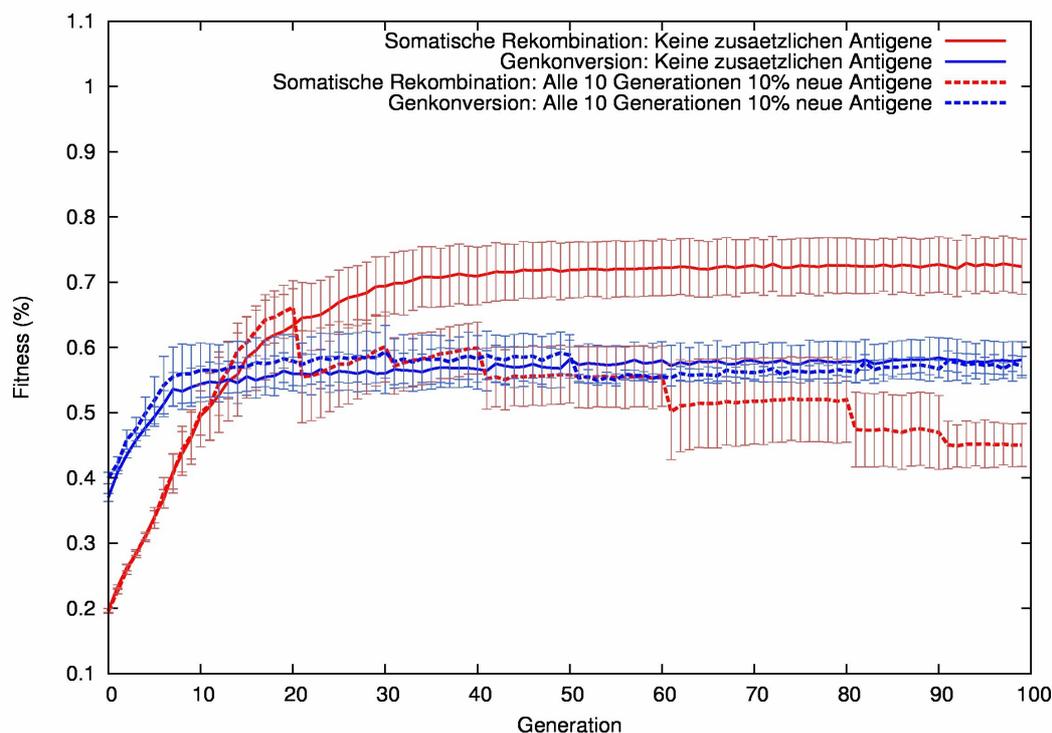


Abbildung 15: Einfluss der Generierung 10% neuer Antigene nach jeweils 10 Generationen auf die Fitness; Parameter: Werte gemäß Tabelle 1; Nach jeweils 10 Generationen wird ein neues Antigen erzeugt; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Ebenso wie bei der vollständigen Neugenerierung der Antigene wird die Genkonversion weniger von neuen Antigenen beeinflusst als die Somatische Rekombination. Bei dieser ist ein stufenförmiger Verlauf der maximalen Fitness nach der 20. Generation zu beobachten. Dabei sinkt

die Fitness jeweils, wenn neue Antigene erzeugt worden sind. Diese kann sich im Anschluss daran allerdings wieder leicht steigern. Das Vermögen sich zu steigern, nimmt mit zunehmender Generationszahl ab. Es ist erwähnenswert, dass die erste Erzeugung neuer Antigene einen positiven Einfluss auf die Fitness hat. Bei der Genkonversion hingegen ändert sich die Fitness im Signifikanzbereich nicht und bleibt damit konstant.

Da jede neue Generation neuen Antigenen ausgeliefert ist, wird kontinuierlich in jeder Generation 1 neues Antigen erzeugt. Es handelt sich dabei um eine Extremsituation.

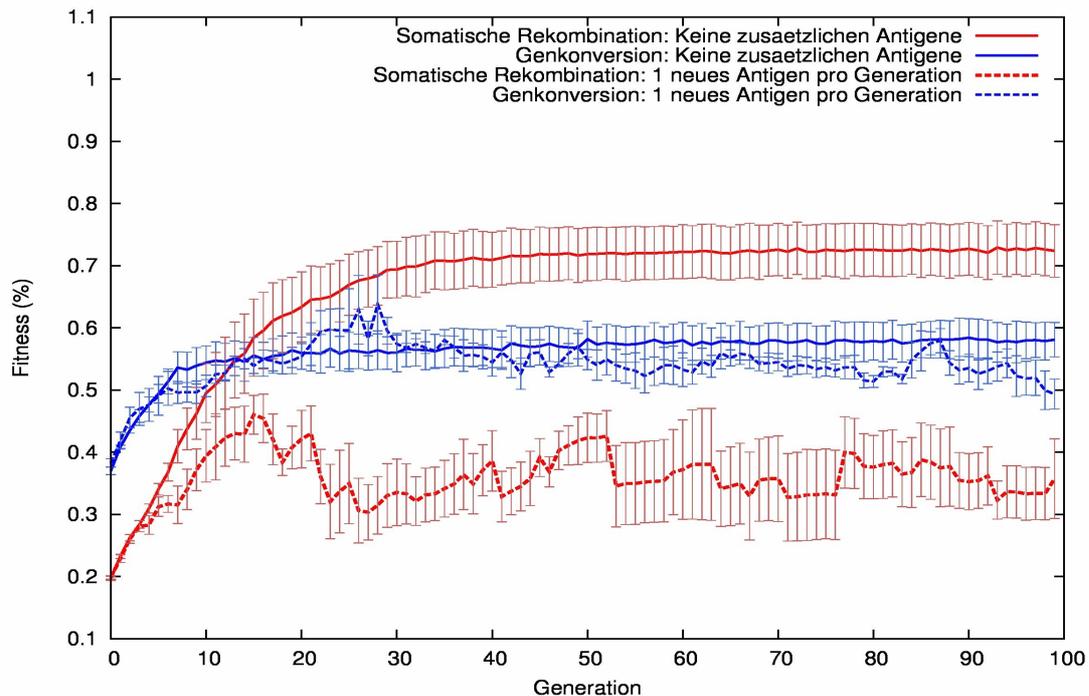


Abbildung 16: Einfluss der Generierung neuer Antigene in jeder Generation auf die Fitness; Parameter: Werte gemäß Tabelle 1; Nach jeder Generation wird ein neues Antigen erzeugt; Fitnesswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Dieser Versuch zeigt, dass eine ständige Veränderung der äußeren Umgebung sowohl einen fitnesssteigernden als auch fitnesssenkenden Einfluss auf die Populationen haben kann. Entsprechend der Ergebnisse aus vorherigen umweltverändernden Versuchen nimmt die Genkonversion allgemein höhere Fitnesswerte an. Bei der Somatischen Rekombination findet bis zur ungefähr 15. Generation nur Fitnesssteigerung statt und erreicht dort ihr Maximum. Danach erreicht sie dieses nicht mehr und zeigt kontinuierliche Schwankungen der Fitness auf, wobei sich zwischenzeitlich Plateaus mit gleich bleibender Fitness ausbilden. Die Genkonversion zeigt geringere Schwankungen. Zudem finden sich hier keine ausgeprägten Plateaus, sondern neben einer eher konstanten Fitness einige kurzzeitige Fitnessänderungen.

5.2.2 Entwicklung der Genvielfalt über Generationen

Um eine noch präzisere Aussage über die beiden Methoden bzw. deren Unterschiede treffen zu können, wird eine neue Methode implementiert, die die Genvielfalt bestimmt. Dazu werden alle Gene von Individuen einer Generation miteinander verglichen und eine prozentuale Übereinstimmung berechnet.

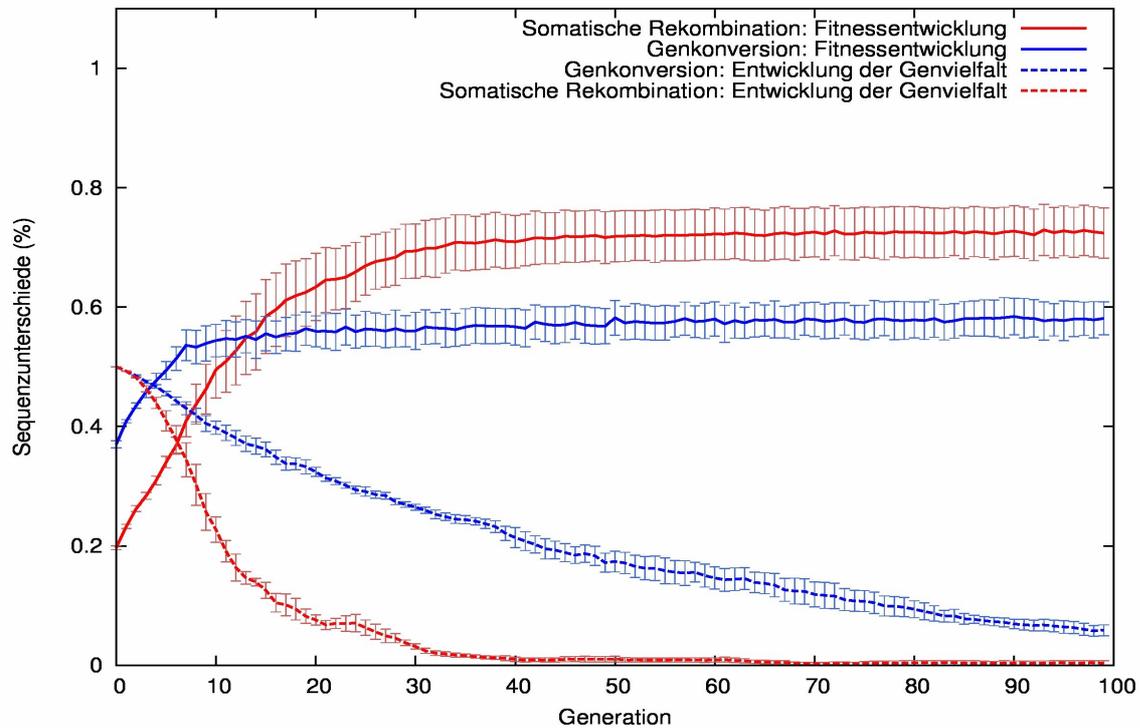


Abbildung 17: Veränderung der Genvielfalt über Generationen; Parameter: Werte gemäß Tabelle 1; Genvielfaltswerte und Standardfehler beruhen auf jeweils vier Simulationen unter identischen Bedingungen

Die Entwicklung der Genvielfalt zeigt deutliche Unterschiede der beiden Mechanismen auf. Bei der Somatischen Rekombination ist die Vielfalt bereits nach 40 Generationen auf einen Wert nahe '0' gefallen. Bei der Genkonversion sinkt die Vielfalt langsamer.

6. Diskussion

Die durchgeführten Versuche machen eine Diskussion über die Evolutionsstrategien des Immunsystems und Unterschiede der Mechanismen in diesem Kapitel möglich. An den Simulationen unter (erweiterten) Standardbedingungen zeigen sich prinzipielle Eigenschaften:

Die Genkonversion beruht auf einem komplexen Zusammenspiel zufälliger Mechanismen, wodurch eine sehr große Antikörpervielfalt generiert werden kann. Dies führt zu hohen Ausgangsfitnesswerten. Vergleichsweise geringer sind ausnahmslos die Fitnesswerte der Startpopulationen bei der Somatischen Rekombination. Die Vielfalt ist durch die Struktur der zu verknüpfenden V- und J-Gene und deren Kombinationsmöglichkeiten begrenzt. Die Ausgangsfitness beider Mechanismen ändert sich bei der Veränderung der Populationsgröße (5.1.2) nicht, bei der Variation der Antikörperanzahl (5.1.3), Antigenanzahl (5.1.4) und der Antigenlänge (5.1.5) hingegen schon. Die Ausgangsfitness beruht ausschließlich auf der Einzelfitness der Individuen und ist unabhängig von der allgemeinen genetischen Vielfalt in der Population, da noch kein Austausch von genetischem Material stattgefunden hat. Sowohl das Verhältnis von Antigen- zu Antikörperanzahl als auch deren Längenverhältnisse nehmen Einfluss auf die Fitness eines Individuums.

Bei der Fitnessentwicklung über die Generationen hinweg ist vor allem die prozentuale Fitnesssteigerung zu betrachten. Bei der Somatischen Rekombination kann sich die Fitness erheblich mehr steigern als bei der Genkonversion. Da bei der Somatischen Rekombination die Gene nicht durch zufällige Mechanismen verändert werden, stehen die Gene und Antikörper in einem direkten Zusammenhang. Dies trifft für die Genkonversion nicht zu, da auf Basis vieler zufällig bestimmter Faktoren zunächst die funktionellen Gene modifiziert und dann die Antikörper gebildet werden. Je unmittelbarer die Gene und Antikörper im Zusammenhang stehen, desto gezielter findet Selektion der für die Antigenanpassung benötigten Gene statt. Dies zeigt sich ebenfalls an den in 5.1.6 durchgeführten Versuchen zur Untersuchung des Einflusses der Anzahl an Genkonversionen: Je geringer das funktionelle Gen verändert wird, desto größer ist die Fitnesssteigerung.

An den erweiterten Versuchen zeigen sich daraus resultierende Konsequenzen. Werden nach einigen Generationen teilweise (5.2.1) oder vollständig (5.1.7) neue Antigene erzeugt, hat dies in Abhängigkeit des verwendeten Mechanismus einen unterschiedlichen Effekt auf die Individuen. Genkonversion betreibende Individuen zeigen nahezu unveränderte Fitnesswerte auf. Bei Individuen, die Somatische Rekombination betreiben, kommt es hingegen zu einem erheblichen Fitnessseinbruch. Eine erneute Fitnesssteigerung ist in größerem Umfang möglich, wenn die Neugeneration zu einem früheren Zeitpunkt stattfindet. Dies beruht auf den unterschiedlichen Strategien

der Mechanismen. Der Somatischen Rekombination gelingt recht präzise Anpassung an die Antigene. Dabei werden alle nicht benötigten Gene aussortiert und die Frequenz der für die Antigenerkennung wichtigen Gene erhöht. Ergebnis dieses Prinzips ist die Verarmung der Vielfalt (5.2.2). Je später die Individuen mit neuen Antigenen konfrontiert werden, desto begrenzter ist deren genetische Vielfalt und umso geringer die Möglichkeit, sich gut an die Antigene anzupassen. Da die Genkonversion sich auf Basis ihres Grundprinzips nicht präzise an die Antigene anpassen kann, bleibt die Vielfalt über mehr Generationen erhalten (5.2.2) und es kann immer wieder auf neue Antigene reagiert werden. Somatische Rekombination führt demnach zu einer exakten Antigenanpassung, Genkonversion hingegen erhält die Vielfalt.

Die Genkonversion ist im Gegensatz zu der Somatischen Rekombination als vollständiger Mechanismus untersucht worden. Da Genkonversion bei wechselnder Umgebung zu höheren Überlebensraten führt, kann man von ihrem Grundprinzip auf die allgemeine Überlebensstrategie des Immunsystems schließen: Verhinderung der Anpassung an die Antigenumgebung. Dies gelingt durch den Einsatz von zufälligen Prinzipien. Zur Untermauerung dieser Vermutung müsste man in einer Simulationserweiterung die junktionale Diversität und die Somatische Hypermuation für die Somatische Rekombination ebenfalls implementieren. Die kombinatorische Diversität alleine reicht jedenfalls nicht aus, um der Strategie nachzukommen.

Auch die Pseudogen- sowie die Antikörpergenerierung könnten bei Simulationserweiterungen verändert werden. Zudem wäre die Betrachtung von leichter und schwerer Kette sinnvoll. Zur Beschleunigung der Versuchsdurchläufe müsste die Methode des Hammingabstandes optimiert werden.

7. Literaturverzeichnis

[Hecker, 2008] N. Hecker. 2008. Primary diversity mechanisms of the antibody synthesis in humans, mice and chickens. *Unveröffentlichte Bachelorarbeit vom Institut für Theoretische Informatik an der Universität in Lübeck*

[Janeway et. al., 2005, 6] C. A. Janeway, P. Travers, M. Walport and M. J. Shlomchik. 2005. Immunology- The immune system in health and diseases. *Garland Science Publishing*

[Arakawa and Buerstedde, 2004] H. Arakawa and J.-M. Buerstedde. 2004. Immunoglobulin Gene Conversion: Insights from bursal B cells and the DT40 cell line *Developmental Dynamics* 229:458-464, 2004

[Campbell and Reece, 2003, 6] N. A. Campbell and J. B. Reece. 2003. Biologie. *Spektrum Verlag*

[Blattman et al., 2002] J. N. Blattman, R. Antia, D. J. D. Sourdive, X. Wang, S. M. Kaech, K. Murali-Krishna, J. D. Altman and R. Ahmed. 2002. Estimating the precursor frequency of naive antigen-specific CD8 T cells. *J. Exp. Med.*, Number 5, 657-664

[Diaz et al., 2001] M. Diaz, M. F. Flajnik and N. Klinman. 2001. Evolution and the molecular basis of somatic hypermutation of antigen receptor genes *Phil. Trans. R. Soc. Lond. B. (2001)* 356, 67-72

[Oprea and Forrest, 1998] M. Oprea and S. Forrest. 1998. Simulated evolution of antibody gene libraries under pathogen selection

[Perelson et al., 1996] A.S. Perelson, R. Hightower and S. Forrest. 1996. Evolution and Somatic learning in v-region genes. *Research in Immunology* 147, 202-208.

[IMGT] ImMunoGenTics-home, IMGT/GENE-DB

<http://www.imgt.org/>

8. Anhang

Tabelle 1: Simulationsparameter

Parameter	Standardwert	Begründung für die Wahl des Standardwertes
Allgemeine Werte für alle Simulationen		
<i>Affinitätsexponent (Affinitätsformel)</i>	40	Auf Basis einer Wahrscheinlichkeitsverteilung: Individuen mit einer höheren Affinität sollen besonders begünstigt werden
<i>Anzahl an Individuen</i>	100	geschätzt
<i>Antikörperanzahl</i>	1000	Eine ausreichende Größe, die sich mit der Speicherkapazität vereinbaren lässt
<i>Antigenanzahl</i>	10	geschätzt
<i>Antigenlänge</i>	400	In Abhängigkeit von der Antikörperlänge angenommen, eigentlich beliebig
<i>Anzahl an Generationen</i>	100	Nach 100 Generationen ist die maximale Fitness erreicht.
<i>Mutationsrate</i>	0,002	[Oprea and Forrest, 1998]
<i>Schwellwert nach Hammingabstandberechnung</i>	0,8	Auf Grund des Affinitätsexponenten
<i>Counter für Antikörpergenerierung</i>	14	Auf Basis von Antikörperlänge
Somatische Rekombination		
<i>V-Genlänge</i>	300	[Janeway et. al., 2005, 6]
<i>J-Genlänge</i>	40	[Janeway et. al., 2005, 6]
<i>Anzahl an V-Genen</i>	40	[Janeway et. al., 2005, 6]
<i>Anzahl an J-Genen</i>	5	[Janeway et. al., 2005, 6]
Genkonversion		
<i>Länge des funktionellen Gens</i>	340	abgeschätzt auf Literaturbasis ([Arakawa and Buerstedde, 2004])
<i>Anzahl an Pseudogenen</i>	25	[Arakawa and Buerstedde, 2004]
<i>Länge der Pseudogene</i>	200	abgeschätzt auf Literaturbasis ([Arakawa and Buerstedde, 2004])