Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 234 (2005) 153–165
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Stochastic Modeling of T cell receptor γ gene rearrangement

Nuno Sepúlveda[a,*], Laurent Boucontet[b], Pablo Pereira[b], Jorge Carneiro[a]

[a]*Instituto Gulbenkian de Ciência, Apartado 14, PT-2781-901 Oeiras, Portugal*
[b]*Institut Pasteur, Unité du Développement des Lymphocytes, rue du Docteur Roux, 25, F-75724 Paris Cedex 15, France*

Received 8 September 2004; received in revised form 11 November 2004; accepted 15 November 2004
Available online 29 December 2004

Abstract

The mechanisms controlling the recombination process of the γ genes that encode the γ chain of the antigen receptor of the $\gamma\delta$ T lymphocytes are unclear. Based on experimental data on the recombination status of the two major TCR γ genes expressed in $V_\gamma 4+$ and $V_\gamma 1+$ thymocytes, we tested the plausibility of three possible rearrangement mechanisms: (1) a time window mechanism according to which the two chromosomes are accessible to the recombination machinery during a defined period of time; (2) a feedback mechanism in which recombination stops shortly after the first in-frame rearrangement event anywhere in both chromosomes; and (3) a feedback mechanism with asynchronous chromosome accessibility, in which there is a first period when only one chromosome is accessible for recombination, followed by a second period when both chromosomes are accessible; shortly after the first in-frame rearrangement event, during any of these two periods, recombination will definitely stop. We model the time window mechanism using a pure probabilistic approach and the two feedback mechanisms using a continuous-time Markov chain formalism. We used maximum likelihood methodology to infer the goodness-of-fit of the models showing evidence for the last model, which best fits the data. Further analysis of this model suggests an evolutionary tradeoff between allelic and isotypic exclusion and the probability that a precursor differentiates into a mature $\gamma\delta$ T lymphocyte.

Keywords: TCR γ gene rearrangement; Markov chains; Time window; Feedback and locus accessibility

1. Introduction

The immune system of vertebrates has a remarkable capacity to recognize and respond to diverse and evolving pathogens. This ability is mainly due to the large diversity of antigen-receptors collectively expressed by B and T lymphocytes (Jerne, 1955; Burnet, 1957). The antigen receptor of B cells is the (Ig) immunoglobulin, which recognizes native proteins, carbohydrates, and lipids, and is composed by two heavy (h) chains and two light chains (either κ or λ). The antigen receptor of T cells (TCR) is a heterodimer made of an α and a β chain in $\alpha\beta$ T cells and γ and δ chains in $\gamma\delta$ T cells. The $\alpha\beta$ TCRs recognize antigenic peptides whereas the ligands of the $\gamma\delta$ TCR are still poorly defined.

The hallmark of Ig and TCR genes is that genes encoding the receptor chains are somatically generated in lymphocyte precursors (Tonegawa, 1983). Different gene segments—V (variable), D (diversity) (for some chains) and J (joining)—are randomly assembled by a process called V(D)J recombination giving rise to the gene that will encode each receptor chain. Since there are many variants of V, D, and J gene segments and some imprecision in the joining a large receptor repertoire can be built by combinatorial assortment.

The basic steps of the V(D)J recombination reaction have been identified. The reaction is initiated by RAG1 and RAG2 proteins (Oettinger et al., 1990) that recognize conserved recombination signal sequences (RSSs) flanking the gene segments, and introduce a double-strand break at the signal sequence boundary

*Corresponding author. Tel.: +351 21 440 7920; fax: +351 21 4407970.
E-mail address: nunosep@igc.gulbenkian.pt (N. Sepúlveda).

0022-5193/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2004.11.018

(Gilfillan et al., 1993). The rearrangement is then completed by ubiquitous DNA repair machinery that joins the gene segments (reviewed in Rooney et al., 2004). While the repair is not finalized, specific enzymes chop off or add nucleotides to the coding sequences in a template independent manner (Komori et al., 1993). Due to the randomness in stitching the gene segments the assembled gene product may not conserve the translation reading frame. Hence, some rearrangements will be in-frame such that the coding sequence can be transcribed and translated to a functional protein, while others will be out-of-frame and therefore non-productive (Coleclough, 1983). The frequency of potential in-frame rearrangements is at maximum of 1/3 (Coleclough, 1983), which may be further decreased if a stop codon is introduced into the reading frame aborting translation of the complete chain (Gu et al., 1991).

Overall control of the V(D)J recombination process is mainly due to the developmental regulation of the expression of RAG genes that are transiently upregulated in lymphocyte precursors (Wilson et al., 1994). Targeting specific loci for recombination is mediated by cis-regulatory elements such as enhancers, promoters and locus control regions that contribute to the local opening of the chromatin structure, and make RSSs accessible to RAG proteins (Yancopoulos and Alt, 1985; Stanhope-Baker et al., 1996; reviewed in Bergman et al., 2003). The efficiency of the recombination is also fine-tuned by the RSSs structure of the different gene segments (reviewed in Feeney et al. (2004)).

The fact that the overwhelming majority of lymphocytes expresses receptor chain proteins from one single allele—a phenomenon known as allelic exclusion—indicates that the mechanisms controlling rearrangement must be very well tuned both individually and collectively. The predominant rationale for the efficiency of allelic exclusion is that rearrangement will be attempted in one gene segment and it will only proceed to the remaining segments if the first attempt was non-productive. Preventing more than one rearrangements is partially mediated by a feedback mechanism whereby the first productive rearrangement shuts down the recombination machinery, likely by repressing the expression of the RAG proteins (Malissen et al., 1992; Aifantis et al., 1997; Uematsu et al., 1988; Alt et al., 1981; Kitamura and Rajewsky, 1992). But efficient allelic exclusion will always require some asynchrony in allele rearrangement that can be ensured in two non-mutually exclusive ways: by forcing the rates of rearrangement to be low such that simultaneous rearrangements are improbable (Cohn and Langman, 1990); and/or by making the two chromosomes become differentially accessible to the recombination machinery (Mostoslavsky et al., 1998, 2001; Khor and Sleckman, 2002). Hitherto, it is unclear how these three processes act in a concerted fashion to ensure that enough mature lymphocytes, which are allelically excluded, are produced from precursors.

This paper focuses on the rearrangement mechanisms of the mouse TCR γ locus, which contains more than one γ gene isotype (Fig. 1). Therefore, $\gamma\delta$ T cell precursors could rearrange productively two isotypes in the same chromosome. The comparison of the extents of allelic and isotypic inclusion in the γ loci (we report elsewhere Boucontet et al. (2004)) can therefore be used to assess the relative importance of chromosome accessibility and feedback mechanisms in controlling rearrangement. Hence, this paper aims to assess how recombination rates, feedback, and loci accessibility come into play to ensure isotypic and allelic exclusion and an adequate yield (i.e. frequency at which a precursor will make at least one functional rearrangement and thus differentiate into a mature cell). Based on experimental data of the recombination status of the two major expressed γ genes in $V_\gamma 4+$ and $V_\gamma 1+$ thymocytes, we investigate three possible rearrangement mechanisms: time window, pure feedback and feedback with asynchronous chromosome accessibility. The first one has both loci accessible in the two chromosomes and the recombination machinery upregulated during a fixed period of time (usually known as time window). The pure feedback mechanism postulates also a time window, but, in addition, when the first in-frame rearrangement event takes place, the recombination machinery is downregulated and/or the loci become non-accessible to this machinery. There is a delay between the in-frame rearrangement and the complete cessation of rearrangements in other loci. The last mechanism extends the previous one, in the sense that the two chromosomes do not become simultaneously accessible. So, there is a first period when only one chromosome is accessible for recombination, and a second period where both chromosomes are accessible.

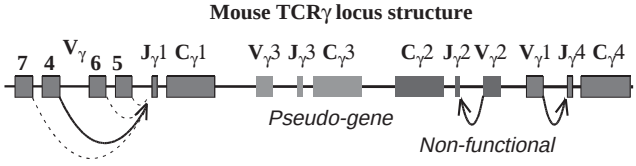

Fig. 1. Structure of the mouse γ gene locus and the rearrangement possibilities. The TCR γ chain locus has four J_γ gene segments. Three of them are associated to a single V_γ segment, whereas the other ($J_\gamma 1$) is associated with four different V_γ segments: $V_\gamma 4$, $V_\gamma 5$, $V_\gamma 6$ and $V_\gamma 7$. The V segments rearrange mainly to the nearest J_γ segment as indicated by the arrows. The solid lines indicate predominant rearrangements in adult mice, in which 80% of mature thymocytes express either $V_\gamma 4J_\gamma 1$ or $V_\gamma 1J_\gamma 4$ genes. The $V_\gamma 5$ and $V_\gamma 6$ rearrange to $J_\gamma 1$ only during fetal life, and $V_\gamma 7$ rearrangement is much less probable than that of $V_\gamma 4$. $V_\gamma 3$ is a pseudogene and $V_\gamma 2$ despite rearranging often in perinatal cells rarely gives rise to functional chains. A specific isotype or allele rearrangement can happen at most once and further rearrangements have to happen elsewhere.