氏　　名	山形哲司
学位（専攻分野）	博士(理学)
学位記番号	総研大甲第334号
学位授与の日付	平成１０年3月２４日
学位授与の要件	生命科学研究科　遺伝学専攻 学位規則第４条第１項該当
学位論文題目	Long-range G+C% mosaic domains in and around the human MHC region; characteristic genome structures and newly found genes in and around the G+C% domain boundaries,with an emphasis on the complete gene structure of GABA receptor B.
論文審査委員	主　査　教　授　五條堀 孝 教　授　西川　建 助教授　齋藤 成也 助教授　城石 俊彦 教　授　猪子　英俊（東海大学）

論文内容の要旨

The human genome is composed of long-range GC% mosaic structures predicted to be related to chromosome bands. DNA replication timing, gene density, and repetitive sequence density have been connected to the chromosome band zones and the long-range GC% mosaic domains. The human MHC spans approximately 4 megabase-pair (Mbp) on the short arm of chromosome 6, and is composed of Mbp-level mosaic domains of GC%, and a boundary of the GC% mosaic domains exists in the junction area between MHC classes II and III. DNA replication timing during S phase is correlated cytogenetically with chromosome band zones, and thus the band boundaries have been predicted to contain a switch point for DNA replication timing. One aim of this thesis study was to determine the precise DNA replication timing for MHC classes II and III, focusing on the junction area. He showed that the replication timing changes precisely in the GC% boundary region with a 2-hour difference of the timing during S phase, supporting the prediction that the GC% boundary is a chromosome band boundary. It was supposed that the replication fork movement terminates (pauses) or significantly slows down in the switch region, which contains dense *Alu* clusters and a long polypurine / polypyrimidine tract with the triple-helix forming potential.

The second aim was to clarify the global GC% distribution in and around the entire 4 Mbp of the human MHC region. He conducted a long-range chromosome walk of about 2.1 Mbp, and found the MHC and its surrounding region to be composed of five long-range GC% mosaic domains disclosing three new GC% mosaic boundaries. One boundary corresponds to the junction between the MHC class II and the centromeric non-MHC region, the second does to the junction between the MHC classes III and I, and the third does to the junction between the class I and the telomeric non-MHC region. It was thus shown that the GC% boundaries correspond to boundaries of functional domains of the respective genome region. In the case of the junction between the MHC class I and the telomeric non-MHC, the non-MHC region was evidently AT-rich, and the telomeric probe was located on 6p22.1 (G/Q) but close to 6p21.3 (R) by a standard fluorescence *in situ* hybridization (FISH) onto prometaphase chromosomes. Since the main body of the MHC region is present on 6p21.3, the walked region was shown to contain the boundary between 6p21.3 and 6p22.1. Characteristic structures in the GC% transition regions were disclosed.

The third aim was to clarify the distribution of genes in the walked genome region, in order to understand the chromosome bands and GC% mosaic structures on functional and evolutionary views. First, he determined sequences of the terminal portions of the cloned fragments of individual cosmids which cover the telomeric portion of the MHC class I and the adjacent non-MHC. Because of the high density of the cosmids covering the individual regions,

portions of each of various genes were found in a series of the consecutive cosmids. Nine genes correspond to those being previously mapped in this region, and nine genes were newly identified in this study; SMT3B-like gene, GABAB receptor gene, mas-related gene, RASH-like gene, TRE17-like gene and MEA11-like gene, in the direction from the class I to the non-MHC, and three olfactory genes. Sequence similarity of some of these genes with those on chromosomes 7, 11, and 17, was observed. This set of chromosomes is distinct from the three chromosomes 1, 9 and 19, on which a wide range of genes with sequence similarities with those on 6p21.3 were present, showing that a possible boundary of the genome multiplication during evolution and/or of the genome rearrangement after the multiplication is located near the telomeric edge of the MHC class I. Elucidation of this type of boundaries and their correlation with boundaries of GC% mosaic domains and of DNA replication timings, should give knowledge of evolutionary processes and mechanisms to establish the present-day human and mammalian genomes.

Among the newly identified genes in this region, he thought that GABR-B is the most interesting gene to be studied and determined its complete genomic sequence. The gamma-aminobutyric acid (GABA) is the most abundant and widely distributed inhibitory neurotransmitter present in the central nervous system. Two mRNA forms of the rat metabotropic gamma-aminobutyric acid receptor (GABAB receptor) have been characterized and predicted to be produced by alternative splicing. He determined a complete sequence of the human GABAB receptor gene. The pairwise alignment of this genomic sequence with that of the known rat GABAB receptor cDNAs showed that the human gene spans 31 kilobases and is composed of 23 exons. The sequence was the first example of the genomic sequence for the GABAB receptor, and thus the first study to characterize the structure for regulating transcription. The two isoforms of the GABAB receptor were found to be generated by alternative usage of promoters, rather than by alternative splicing. In both promoter regions, CpG islands and several potential transcription regulatory sites were identified. A cAMP response element (CRE) was present only in the promoter for the shorter isoform. Based on the fact that the GABAB receptor negatively regulates the CRE binding protein (CREB) - mediated transcriptions in the central nervous system, he proposed that the activation of GABAB receptor can differentiate the relative expression of its two isoforms through alternative usage of the promoters with and without a CRE, via the CRE- and CREB-mediated regulatory system. This may modulate synaptic transmission. An informative polymorphic marker (CA)n was found 1 kb upstream of the 5'-UTR of the GABAB receptor gene, and used to analyze 36 different HLA haplotypes.

論文の審査結果の要旨

ヒトゲノムには、数百Kbpから数Mbpの広範囲なDNA領域にわたってGC%が非常に高い部分がところどころに存在し、いわばヒトゲノムはGC%の高い領域とそうでない領域のモザイク構造になっている。このモザイク構造の機能的な意義や進化的起源については多くの議論のあるところであり、未知に残された問題が多い。山形君は、GC%の高い領域と低い領域の境界の例がヒト染色体6番の短腕側に位置する約4Mbp領域の主要組織連合性複合体(MHC)遺伝子群の中にあることに注目した。この境界領域はMHCクラスIIとクラスIIIの間をつなぐ領域に存在することがわかっている。DNAのCG含量がS期におけるDNA複製時期と関係するとすれば、DNA複製時期はGC%モザイクの境界をはさむMHCクラスIIとクラスIIIとで異なることが予想されるので、それらの領域におけるDNA複製時期を実験的に測定した。その結果、この境界をはさむGC%の高い領域とそうでない領域の間には2時間もの差が存在することを明らかにした。このことから、この境界においては、DNAの複製forkの動きが一時的に停止するか少なくとも顕著に減速されることを明らかにした。さらに、4MbpにわたるヒトMHC領域全体について、YACやコスミドを用いて広範囲な染色体walkingの実験を行い、GC含量分布を測定した。その結果、GC%モザイク領域の境界がMHCクラスIIとセントロメア側の非MHC領域の境界、MHCクラスIIIとクラスIの間、MHCクラスIとテロメア側の非MHC領域の境界にも存在することを明らかにした。また、walkingを行ったゲノム領域において遺伝子の存在密度を調べる過程で、9つの新しい遺伝子を発見した。その中でも、GABR-BはGABA receptorの一つと考えられ、その遺伝子の完全配列を決定した。これは、GABA-B receptorのゲノム配列の最初の例となった。また、この遺伝子配列と他の相同な遺伝子配列を比較検討し、興味深い転写制御のメカニズムを提案した。

これらの研究は、ヒトゲノムの機能や構造及びその進化的意義を考察する上で、新しい知見を多く含んでおり、学術的に重要な貢献と考えられる。また、本研究は分子遺伝学やゲノム進化学の基礎的素養を修得し駆使した研究で、学位論文にふさわしく先駆的である。よって、学位を授与するにふさわしい論文と判定した。