

FIG. 8. Changes in staining of germinal center markers in normal or SRBC-immunized *Cmah*-null mice. (A) Histochemical analyses of spleen sections without immunization. Spleen sections from wild-type and *Cmah*-null mice were costained with FITC-conjugated GL7 and biotin-conjugated PNA visualized by R-PE-conjugated streptavidin. (B) Histochemical analyses of spleen sections after T-dependent immunization. Wild-type and *Cmah*-null mice were immunized with SRBC, and the spleens were removed 8 days after immunization. The frozen spleen sections were costained with FITC-conjugated GL7 and biotin-conjugated PNA followed by R-PE-conjugated streptavidin. Arrows indicate germinal centers. Genotypes are indicated as follows: +/+, wild type; -/-, *Cmah* null.

alteration of surface-sialylated glycans and that this alteration of Sia species from Neu5Gc to Neu5Ac can be probed with GL7. This is the first report regarding the epitope identification of GL7, which is routinely used to stain germinal center B cells in mice. We demonstrated that the GL7 epitope is the Neu5Ac α 2-6LacNAc-containing N-glycan, which is prominently expressed in activated B cells upon the repression of *Cmah*. Gain of GL7 epitope expression coincided with the loss of optimal ligand expression of CD22 in germinal center B cells, presumably centrocytes. Considering the rather strong degree of GL7 positivity in germinal center B cells in comparison with in vitro stimulated B-cell blasts, the degree of *Cmah* reduction might have been severe in these cells. In general, it is thought that Neu5Gc is easy to accumulate but difficult to turn over in cells. This is attributable to the one-way direction of the metabolic pathway; Neu5Gc is biosynthesized by *Cmah* from Neu5Ac (24, 36, 54), whereas no conversion activity was found to biosynthesize Neu5Ac from Neu5Gc. Therefore, the reduction of Neu5Gc found in the GL7-enriched germinal center cells is remarkable. Such rapid clearance of Neu5Gc could be attributable to several characteristics of germinal center cells. Most importantly, as shown in Fig. 3D, these cells repressed *Cmah*, the enzyme responsible for the de novo biosynthesis of Neu5Gc. Moreover, because lymphocytes are small cells with limited cytosolic space, the cytosolic pool of Sia in these cells is likely limited and easily turned over. In addition, centrocytes undergo extremely fast cell cycles (66), which probably leads to rapid passive dilution of the cytosolic pool in these cells. At the same time, new protein synthesis should be a primary event that happens in germinal center B cells, as shown by cDNA microarray analysis (51). The transcriptional repression of *Cmah*, together with these features of germinal center cells, could contribute to the efficient conversion of the major Sia species from Neu5Gc to Neu5Ac.

Negative regulation of B-cell activation by *Cmah* and its product, Neu5Gc. To clarify the biological role of Neu5Gc in

vivo, we disrupted the *Cmah* gene in mice and examined their B-cell activation phenotypes. *Cmah*-null mice showed a hyperreactive B-cell phenotype to T-independent stimulation. In contrast, the T-dependent immunization response was similar to that in wild-type mice. This is consistent with the findings that *Cmah* expression is severely repressed in the germinal centers of wild-type spleen upon T-dependent immunization and that *Cmah*-null mice could develop follicles stained with PNA, another marker for germinal centers. Forced expression of *Cmah* caused repression of the proliferative response of *Cmah*-null B cells, indicating that Neu5Gc-containing sialoglycan functions to suppress B-cell reactivity though the mechanism is still unknown. This suppression via Neu5Gc-containing sialoglycan appears to be canceled by *Cmah* repression in germinal center B cells that are "activation committed" or "activation competent." The hyperreactive B-cell phenotypes observed in *Cmah*-null B cells could mirror differences in cellular reactivity between germinal center and nongerminal center B cells, as indicated by differential cell surface expression of the GL7 epitope (5).

Possible change in sialoglycan-receptor interaction in *Cmah*-null mice. As *Cmah* disruption results in a single oxygen atom change in these mice, it is expected that this mutation leaves both the Sia amount and Sia linkage intact in terms of sialoglycans, which could change the stability or turnover of the proteins modified with Sia (14). Although only limited information is available, sialyltransferases that biosynthesize sialylated glycans in the Golgi apparatus do not show strong preferences for CMP-Neu5Ac or CMP-Neu5Gc as substrates (59). When we probed linkage-specific protein sialylation by using α 2,6-linked Sia-binding plant lectins such as *Sambucus nigra* agglutinin, we did not observe a change (data not shown). Thus, the molecular event affected in *Cmah*-null mice is likely to be lectin recognition of a single oxygen atom on sialoglycans expressed on the cell surface, although a single responsible lectin may not explain the phenotype. One of the candidate lectins as the receptor of sialoglycans is the Siglec family (9, 12, 62), though a yet-to-be-characterized Sia-binding molecule could be affected.

When ligand expression for Siglecs was detected using Siglec-Fc probes, *Cmah*-null mice lost optimal ligand expression for CD22 (Siglec-2). The ligand function of CD22 in a mouse model has been addressed in two different ways. One study was done using *St6gal1*-knockout mice (20), and another study analyzed gene-targeted mice expressing mutant CD22 molecules that do not interact with ligands (43). The phenotypes found in *Cmah*-null mice are considerably different from these two previous studies; therefore, *Cmah*-null phenotypes might be caused by the combination of loss/gain of a Sia-mediated interaction. Additional studies using a combination of various knockout strains related to sialoglycan recognition are required to address such possibilities.

Apart from the phenotypic contribution of CD22 to the assays in the present study, CD22 ligand expression is not static but is, instead, a regulated event during in vivo B-cell activation. We showed that mCD22-Fc probe staining was down-regulated in germinal centers. Moreover, it was reported that in vitro activated human B cells unmask CD22 from a *cis*-ligand (45). Thus, the regulation of CD22 ligand expression

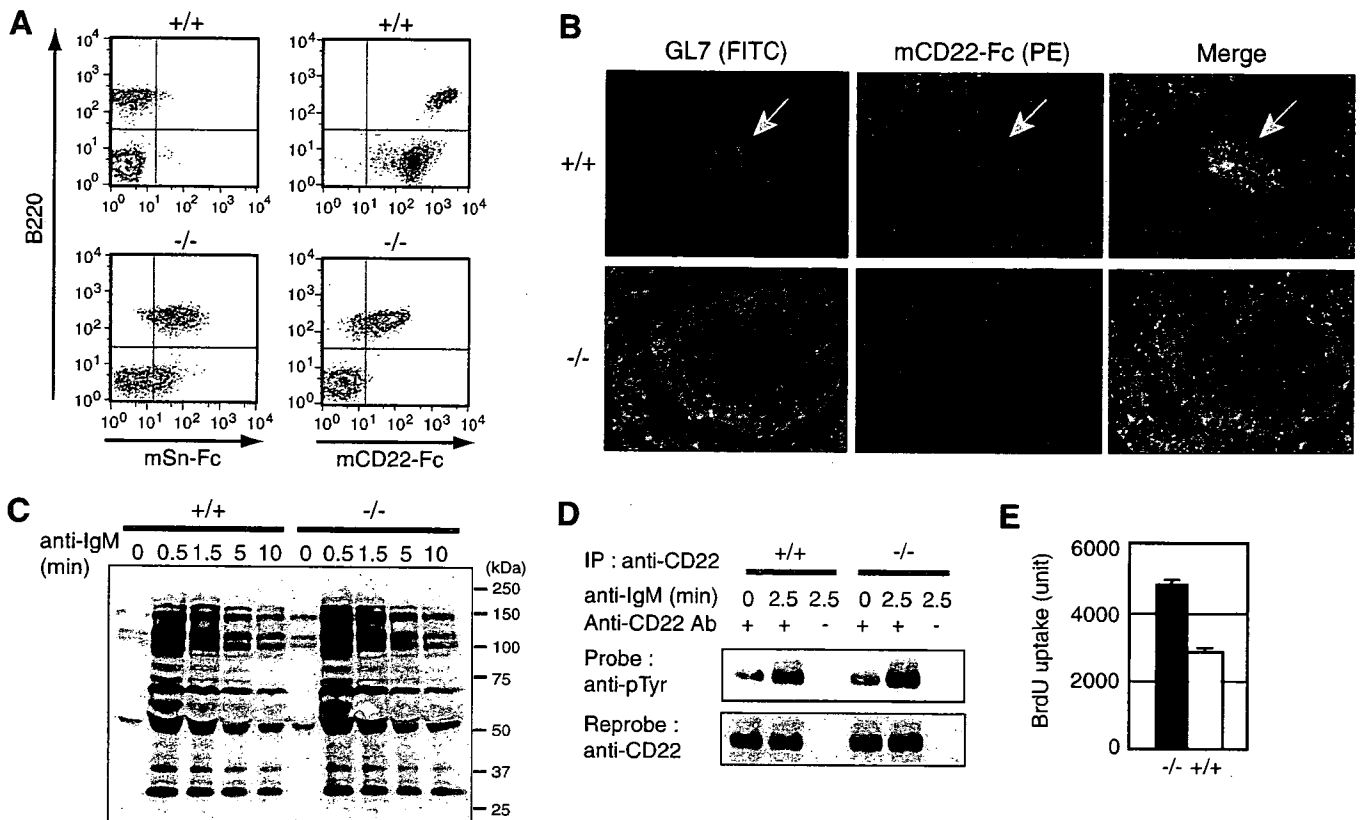


FIG. 9. Loss of optimal CD22 ligand and normal immediate response upon BCR cross-linking in *Cmah*-null mice. (A) Loss of optimal ligand for CD22 in *Cmah*-null mice. The expression of surface ligands for sialoadhesin and CD22 was detected by flow cytometry. Splenocytes from wild-type and *Cmah*-null mice were costained with FITC-conjugated anti-B220 and mSn/mCD22-Fc precomplexed with R-PE-conjugated anti-human IgG. Wild-type B cells were strongly stained with mCD22-Fc. In contrast, the level of mCD22-Fc staining showed a marked decrease in *Cmah*-null mice. The weak signal found on *Cmah*-null splenocytes was detected only with the chimeric probe mCD22-Fc prepared from Lec2 cell culture medium and not with the probe prepared from COS7 cells, possibly because of the autosialylation. (B) Histochemical analyses of CD22 ligand expression in spleen sections. Spleen sections from wild-type and *Cmah*-null mice 8 days after SRBC immunization were costained with FITC-conjugated GL7 and mCD22-Fc precomplexed with R-PE-conjugated anti-human IgG. Arrows indicate germinal centers. (C) Overall tyrosine phosphorylation upon anti-IgM stimulation. Splenic B cells from wild-type and *Cmah*-null mice were stimulated with the F(ab')₂ fragment of anti-mouse IgM (μ chain) for the indicated times. Whole-cell lysates were subjected to immunoblotting with antiphosphotyrosine antibody (PT-66). (D) Phosphorylation of CD22. Splenic B cells were stimulated with the F(ab')₂ fragment of anti-mouse IgM (μ chain) for the indicated times. The cell lysates were subjected to immunoprecipitation with anti-CD22 antibody (Cy34.1). The precipitated proteins were immunoblotted with antiphosphotyrosine (pTyr) antibody (PT-66) and then reprobated with anti-CD22 polyclonal antibody. (E) In vitro hyperproliferation response of *Cmah*-null B cells to calcium signaling. Splenic B cells were assessed for proliferation using tetradecanoyl phorbol acetate (10 ng/ml) plus ionomycin (5 μ g/ml) as stimulating reagents. The proliferation assay was performed as described in the legend of Fig. 6C. The open column represents the mean proliferation of wild-type B cells, and the filled column represents the mean proliferation of *Cmah*-null B cells. The bars represent the standard errors of the mean for triplicate cultures. +/+, wild type; -/-, *Cmah* null; IP, immunoprecipitation.

could be an important event to modulate B-cell activation in vivo.

Loss of Neu5Gc in relation to human deficiency for the CMAH gene. *Homo sapiens* is the sole mammalian species that lacks Neu5Gc expression throughout the body; indeed, Neu5Gc is antigenic to humans (31). This is a striking difference between humans and chimpanzees, which express Neu5Gc as the major species of Sia throughout their bodies. Recently, it was shown that, unlike gene expression in the extant great apes, the *CMAH* gene is inactivated in humans (6, 22). Here, we demonstrated that *Cmah* is the sole enzyme responsible for the production of Neu5Gc in cells since our mouse model reproduced the human-like deficiency in Neu5Gc biosynthesis. This result confirmed that a genetic mu-

tation in the human lineage caused the lack of Neu5Gc in humans.

Sia is commonly used in the host recognition system of microbes, and human-specific microbes are reported to recognize epitope(s) containing Neu5Ac on human cells. The mouse described here is thus the first mammalian line that could be used as an animal model system to assess Sia-targeted human infectious diseases (15).

ACKNOWLEDGMENTS

We thank Motomi Osato and Yoshiaki Itoh for the blood chemistry and blood counting experiments. We also thank Ajit Varki, Takeshi Tsubata, and Reiji Kannagi for helpful discussions during the preparation of the manuscript.

This work was supported by CREST, Japanese Science and Technology; a grant-in-aid program from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and RIKEN.

REFERENCES

- Alizadeh, A. A., M. B. Eisen, R. E. Davis, C. Ma, I. S. Lossos, A. Rosenwald, J. C. Boldrick, H. Sabet, T. Tran, X. Yu, J. I. Powell, L. Yang, G. E. Marti, T. Moore, J. Hudson, Jr., L. Lu, D. B. Lewis, R. Tibshirani, G. Sherlock, W. C. Chan, T. C. Greiner, D. D. Weisenburger, J. O. Armitage, R. Warnke, R. Levy, W. Wilson, M. R. Grever, J. C. Byrd, D. Botstein, P. O. Brown, and L. M. Staudt. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–511.
- Baum, L. G., K. Derbin, N. L. Perillo, T. Wu, M. Pang, and C. Uittenbogaart. 1996. Characterization of terminal sialic acid linkages on human thymocytes. Correlation between lectin-binding phenotype and sialyltransferase expression. *J. Biol. Chem.* 271:10793–10799.
- Blixt, O., B. E. Collins, I. M. Van Den Nieuwenhof, P. R. Crocker, and J. C. Paulson. 2003. Sialoside specificity of the Siglec family assessed using novel multivalent probes: identification of potent inhibitors of myelin associated glycoproteins. *J. Biol. Chem.* 278:31007–31019.
- Buhl, A. M., and J. C. Cambier. 1997. Co-receptor and accessory regulation of B-cell antigen receptor signal transduction. *Immunol. Rev.* 160:127–138.
- Cervenak, L., A. Magyar, R. Boja, and G. Laszlo. 2001. Differential expression of GL7 activation antigen on bone marrow B cell subpopulations and peripheral B cells. *Immunol. Lett.* 78:89–96.
- Chou, H. H., H. Takematsu, S. Diaz, J. Iber, E. Nickerson, K. L. Wright, E. A. Muchmore, D. L. Nelson, S. T. Warren, and A. Varki. 1998. A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. *Proc. Natl. Acad. Sci. USA* 95:11751–11756.
- Clark, E. A. 1993. CD22, a B cell-specific receptor, mediates adhesion and signal transduction. *J. Immunol.* 150:4715–4718.
- Coico, R. F., B. S. Bhogal, and G. J. Thorbecke. 1983. Relationship of germinal centers in lymphoid tissue to immunologic memory. VI. Transfer of B cell memory with lymph node cells fractionated according to their receptors for peanut agglutinin. *J. Immunol.* 131:2254–2257.
- Crocker, P. R. 2002. Siglecs: sialic-acid-binding immunoglobulin-like lectins in cell-cell interactions and signalling. *Curr. Opin. Struct. Biol.* 12:609–615.
- Crocker, P. R., S. Kelm, C. Dubois, B. Martin, A. S. McWilliam, D. M. Shotton, J. C. Paulson, and S. Gordon. 1991. Purification and properties of sialoadhesin, a sialic acid-binding receptor of murine tissue macrophages. *EMBO J.* 10:1661–1669.
- Crocker, P. R., and A. Varki. 2001. Siglecs in the immune system. *Immunology* 103:137–145.
- Crocker, P. R., and A. Varki. 2001. Siglecs, sialic acids and innate immunity. *Trends Immunol.* 22:337–342.
- Cyster, J. G., and C. C. Goodnow. 1997. Tuning antigen receptor signaling by CD22: Integrating cues from antigens and the microenvironment. *Immunity* 6:509–517.
- Ellies, L. G., D. Ditto, G. G. Levy, M. Wahrenbrock, D. Ginsburg, A. Varki, D. T. Le, and J. D. Marth. 2002. Sialyltransferase ST3Gal-IV operates as a dominant modifier of hemostasis by concealing asialoglycoprotein receptor ligands. *Proc. Natl. Acad. Sci. USA* 99:10042–10047.
- Gagneux, P., and A. Varki. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9:747–755.
- Han, H., D. J. Bearss, L. W. Browne, R. Calalupe, R. B. Nagle, and D. D. Von Hoff. 2002. Identification of differentially expressed genes in pancreatic cancer cells using cDNA microarray. *Cancer Res.* 62:2890–2896.
- Han, S., S. R. Dillon, B. Zheng, M. Shimoda, M. S. Schlissel, and G. Kelsoe. 1997. V(D)J recombinase activity in a subset of germinal center B lymphocytes. *Science* 278:301–305.
- Han, S., B. Zheng, D. G. Schatz, E. Spanopoulou, and G. Kelsoe. 1996. Neoteny in lymphocytes: Rag1 and Rag2 expression in germinal center B cells. *Science* 274:2094–2097.
- Hathcock, K. S., C. E. Pucillo, G. Laszlo, L. Lai, and R. J. Hodes. 1995. Analysis of thymic subpopulations expressing the activation antigen GL7: expression, genetics, and function. *J. Immunol.* 155:4575–4581.
- Hennet, T., D. Chui, J. C. Paulson, and J. D. Marth. 1998. Immune regulation by the ST6Gal sialyltransferase. *Proc. Natl. Acad. Sci. USA* 95:4504–4509.
- Hokazono, Y., T. Adachi, M. Wabl, N. Tada, T. Amagasa, and T. Tsubata. 2003. Inhibitory coreceptors activated by antigens but not by anti-Ig heavy chain antibodies install requirement of costimulation through CD40 for survival and proliferation of B cells. *J. Immunol.* 171:1835–1843.
- Irie, A., S. Koyama, Y. Kozutsumi, T. Kawasaki, and A. Suzuki. 1998. The molecular basis for the absence of *N*-glycolylneuraminic acid in humans. *J. Biol. Chem.* 273:15866–15871.
- Itoharu, S., P. Mombaerts, J. Lafaille, J. Iacomini, A. Nelson, A. R. Clarke, M. L. Hooper, A. Farr, and S. Tonegawa. 1993. T cell receptor delta gene mutant mice: independent generation of alpha beta T cells and programmed rearrangements of gamma delta TCR genes. *Cell* 72:337–348.
- Kawano, T., S. Koyama, H. Takematsu, Y. Kozutsumi, H. Kawasaki, S. Kawashima, T. Kawasaki, and A. Suzuki. 1995. Molecular cloning of cytidine monophospho-*N*-acetylneuraminic acid hydroxylase. Regulation of species- and tissue-specific expression of *N*-glycolylneuraminic acid. *J. Biol. Chem.* 270:16458–16463.
- Kawano, T., Y. Kozutsumi, T. Kawasaki, and A. Suzuki. 1994. Biosynthesis of *N*-glycolylneuraminic acid-containing glycoconjugates. Purification and characterization of the key enzyme of the cytidine monophospho-*N*-acetylneuraminic acid hydroxylation system. *J. Biol. Chem.* 269:9024–9029.
- Kelm, S., A. Pelz, R. Schauer, M. T. Filbin, S. Tang, M. E. de Bellard, R. L. Schnaar, J. A. Mahoney, A. Hartnell, P. Bradfield, et al. 1994. Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of sialic acid-dependent adhesion molecules of the immunoglobulin superfamily. *Curr. Biol.* 4:965–972.
- Koyama, S., T. Yamaji, H. Takematsu, T. Kawano, Y. Kozutsumi, A. Suzuki, and T. Kawasaki. 1996. A naturally occurring 46-amino-acid deletion of cytidine monophospho-*N*-acetylneuraminic acid hydroxylase leads to a change in the intracellular distribution of the protein. *Glycoconj. J.* 13:353–358.
- Laszlo, G., K. S. Hathcock, H. B. Dickler, and R. J. Hodes. 1993. Characterization of a novel cell-surface molecule expressed on subpopulations of activated T and B cells. *J. Immunol.* 150:5252–5262.
- Lossos, I. S., A. A. Alizadeh, M. B. Eisen, W. C. Chan, P. O. Brown, D. Botstein, L. M. Staudt, and R. Levy. 2000. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc. Natl. Acad. Sci. USA* 97:10209–10213.
- MacLennan, I. C. 1994. Germinal centers. *Annu. Rev. Immunol.* 12:117–139.
- Martin, M. J., A. Muotri, F. Gage, and A. Varki. 2005. Human embryonic stem cells express an immunogenic nonhuman sialic acid. *Nat. Med.* 11:228–232.
- McHeyzer-Williams, L. J., M. Cool, and M. G. McHeyzer-Williams. 2000. Antigen-specific B cell memory: expression and replenishment of a novel B220⁺ memory B cell compartment. *J. Exp. Med.* 191:1149–1166.
- McHeyzer-Williams, L. J., D. J. Driver, and M. G. McHeyzer-Williams. 2001. Germinal center reaction. *Curr. Opin. Hematol.* 8:52–59.
- Mills, D. M., J. C. Stolpa, and J. C. Cambier. 2004. Cognate B cell signaling via MHC class II: differential regulation of B cell antigen receptor and MHC class II/Ig- α signaling by CD22. *J. Immunol.* 172:195–201.
- Morita, S., T. Kojima, and T. Kitamura. 2000. Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* 7:1063–1066.
- Muchmore, E. A., M. Milewski, A. Varki, and S. Diaz. 1989. Biosynthesis of *N*-glycolylneuraminic acid. The primary site of hydroxylation of *N*-acetylneuraminic acid is the cytosolic sugar nucleotide pool. *J. Biol. Chem.* 264:20216–20223.
- Muramatsu, M., V. S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N. O. Davidson, and T. Honjo. 1999. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274:18470–18476.
- Murasawa, M., S. Okada, S. Obata, M. Hatano, H. Moriya, and T. Tokuhisa. 2002. GL7 defines the cycling stage of pre-B cells in murine bone marrow. *Eur. J. Immunol.* 32:291–298.
- Nitschke, L. 2005. The role of CD22 and other inhibitory co-receptors in B-cell activation. *Curr. Opin. Immunol.* 17:290–297.
- Novorodovskaya, N., M. L. Whitfield, L. S. Basehore, A. Novorodovsky, R. Pesich, J. Usary, M. Karaca, W. K. Wong, O. Aprelikova, M. Fero, C. M. Perou, D. Botstein, and J. Braman. 2004. Universal reference RNA as a standard for microarray experiments. *BMC Genomics* 5:20.
- Pasare, C., and R. Medzhitov. 2005. Control of B-cell responses by Toll-like receptors. *Nature* 438:364–368.
- Paulson, J. C., J. Weinstein, and A. Schauer. 1989. Tissue-specific expression of sialyltransferases. *J. Biol. Chem.* 264:10931–10934.
- Poe, J. C., Y. Fujimoto, M. Hasegawa, K. M. Haas, A. S. Miller, I. G. Sanford, C. B. Bock, M. Fujimoto, and T. F. Tedder. 2004. CD22 regulates B lymphocyte function in vivo through both ligand-dependent and ligand-independent mechanisms. *Nat. Immunol.* 5:1078–1087.
- Powell, L. D., and A. Varki. 1994. The oligosaccharide binding specificities of CD22 beta, a sialic acid-specific lectin of B cells. *J. Biol. Chem.* 269:10628–10636.
- Razi, N., and A. Varki. 1998. Masking and unmasking of the sialic acid-binding lectin activity of CD22 (Siglec-2) on B lymphocytes. *Proc. Natl. Acad. Sci. USA* 95:7469–7474.
- Reichert, R. A., W. M. Gallatin, I. L. Weissman, and E. C. Butcher. 1983. Germinal center B cells lack homing receptors necessary for normal lymphocyte recirculation. *J. Exp. Med.* 157:813–827.
- Schauer, R. 1982. Sialic acids: chemistry, metabolism and function. *Cell biology monographs*, vol. 10. Springer-Verlag, New York, NY.
- Schulte, R. J., M. A. Campbell, W. H. Fischer, and B. M. Sefton. 1992. Tyrosine phosphorylation of CD22 during B cell activation. *Science* 258:1001–1004.
- Schwarzkopf, M., K. P. Knobloch, E. Rohde, S. Hinderlich, N. Wiechens, L. Lucka, I. Horak, W. Reutter, and R. Horstkorte. 2002. Sialylation is essential for early development in mice. *Proc. Natl. Acad. Sci. USA* 99:5267–5270.
- Sgroi, D., A. Varki, S. Braesch-Andersen, and I. Stamenkovic. 1993. CD22,

- a B cell-specific immunoglobulin superfamily member, is a sialic acid-binding lectin. *J. Biol. Chem.* **268**:7011–7018.
51. Shaffer, A. L., A. Rosenwald, E. M. Hurt, J. M. Giltman, L. T. Lam, O. K. Pickeral, and L. M. Staudt. 2001. Signatures of the immune response. *Immunity* **15**:375–385.
 52. Shapiro-Shelef, M., K. I. Lin, L. J. McHeyzer-Williams, J. Liao, M. G. McHeyzer-Williams, and K. Calame. 2003. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity* **19**:607–620.
 53. Shaw, L., and R. Schauer. 1988. The biosynthesis of *N*-glycolylneuraminic acid occurs by hydroxylation of the CMP-glycoside of *N*-acetylneuraminic acid. *Biol. Chem. Hoppe-Seyler* **369**:477–486.
 54. Shaw, L., and R. Schauer. 1989. Detection of CMP-*N*-acetylneuraminic acid hydroxylase activity in fractionated mouse liver. *Biochem. J.* **263**:355–363.
 55. Shih, T. A., E. Meffre, M. Roederer, and M. C. Nussenzweig. 2002. Role of BCR affinity in T cell dependent antibody responses in vivo. *Nat. Immunol.* **3**:570–575.
 56. Sjoberg, E. R., L. D. Powell, A. Klein, and A. Varki. 1994. Natural ligands of the B cell adhesion molecule CD22 beta can be masked by 9-*O*-acetylation of sialic acids. *J. Cell Biol.* **126**:549–562.
 57. Takematsu, H., S. Diaz, A. Stoddart, Y. Zhang, and A. Varki. 1999. Lysosomal and cytosolic sialic acid 9-*O*-acetyltransferase activities can be encoded by one gene via differential usage of a signal peptide-encoding exon at the N terminus. *J. Biol. Chem.* **274**:25623–25631.
 58. Tedder, T. F., J. Tuscano, S. Sato, and J. H. Kehrl. 1997. CD22, a B lymphocyte-specific adhesion molecule that regulates antigen receptor signaling. *Annu. Rev. Immunol.* **15**:481–504.
 59. Tsuji, S. 1996. Molecular cloning and functional analysis of sialyltransferases. *J. Biochem. (Tokyo)* **120**:1–13.
 60. Varki, A. 1992. Diversity in the sialic acids. *Glycobiology* **2**:25–40. (Erratum, **2**:168.)
 61. Varki, A. 1997. Sialic acids as ligands in recognition phenomena. *FASEB J.* **11**:248–255.
 62. Varki, A., and T. Angata. 2006. Siglecs—the major subfamily of I-type lectins. *Glycobiology* **16**:1R–27R.
 63. Wakabayashi, C., T. Adachi, J. Wienands, and T. Tsubata. 2002. A distinct signaling pathway used by the IgG-containing B cell antigen receptor. *Science* **298**:2392–2395.
 64. Wuensch, S. A., R. Y. Huang, J. Ewing, X. Liang, and J. T. Lau. 2000. Murine B cell differentiation is accompanied by programmed expression of multiple novel β -galactoside α 2, 6-sialyltransferase mRNA forms. *Glycobiology* **10**:67–75.
 65. Yamaji, T., T. Teranishi, M. S. Alphey, P. R. Crocker, and Y. Hashimoto. 2002. A small region of the natural killer cell receptor, Siglec-7, is responsible for its preferred binding to α 2,8-disialyl and branched α 2,6-sialyl residues. A comparison with Siglec-9. *J. Biol. Chem.* **277**:6324–6332.
 66. Zhang, J., I. C. MacLennan, Y. J. Liu, and P. J. Lane. 1988. Is rapid proliferation in B centroblasts linked to somatic mutation in memory B cell clones? *Immunol. Lett.* **18**:297–299.

第4回(平成18年度)薬学研究ビジョン部会賞受賞者(1)

ケミカル・バイオ情報に基づく創薬インフォマティクス研究

奥野 恭史 京都大学大学院薬学研究科 統合薬学フロンティア教育センター

1. はじめに

ヒトゲノムが解読された今日、莫大なゲノム情報から創薬への手がかりを発見すること、すなわち「ゲノム創薬」に大きな期待が寄せられている。ゲノム創薬は、ゲノム情報を出発点とし創薬の標的遺伝子探索からリード化合物探索を経て臨床段階に至る広範で高度に専門化した複合領域であり、その実践にはこれらの複合領域の橋渡しを実現する統合的なインフォマティクス基盤(創薬インフォマティクス)が必須となる。我々は、創薬インフォマティクスという新たな研究分野の創成に向け、バイオ情報を扱うバイオインフォマティクスとケミカル情報を扱うケモインフォマティクスの独立に発展してきた2つの情報科学分野の統合を図り、バイオ情報とケミカル情報の両者を同時に統合的にマイニングする新しい情報技術の開発に着手している。さらに本研究は、現在、国内外で注目されているケミカルゲノミクス・ケミカルバイオロジーのための有力な情報基盤ともなり得るものと考えられる。

2. 化合物の宇宙探索(ケミカル空間の探索)

2004年12月のNature誌において、Chemical Space特集号が発表された(1)。そこでは、化合物の種類は1060個を越える天文学的なバリエーションを有しており、化合物空間を探索することは宇宙探索と同様に壮大な課題であることが提示されている。このことは医薬品の候補化合物となり得る新規な活性化合物を見つけ出すことが如何に困難でセレンディップなことであるかを示唆するものである。

これらケミカル空間の探索の基礎研究としてケミカルゲノミクス・ケミカルバイオロジー研究が近年注目されている。ケミカルゲノミクスでは、その命題として「莫大な数の化合物と生体系(タンパク質や細胞など)との相互作用を包括的に明らかにすること」が挙げられている。実際、米国では、ケミカルゲノミクスプロジェクトを掲げ、数百万もの膨大な化合物に関する情報を収集し、

有用化合物の探索に国策として取り組んでいる。

しかしながら、広大な化合物空間から生物活性を有する化合物を探し当てる化合物探索には、天文学的な数量に対応できる新たなインフォマティクス技術とハイスループット技術の研究開発が必須である。そこで、我々は、莫大な化合物群とタンパク質群との相互作用様式をゲノムスケールで解析することを目的とした情報学的技術、すなわちケミカルゲノミクスのためのインフォマティクス技術の研究開発を行っている。

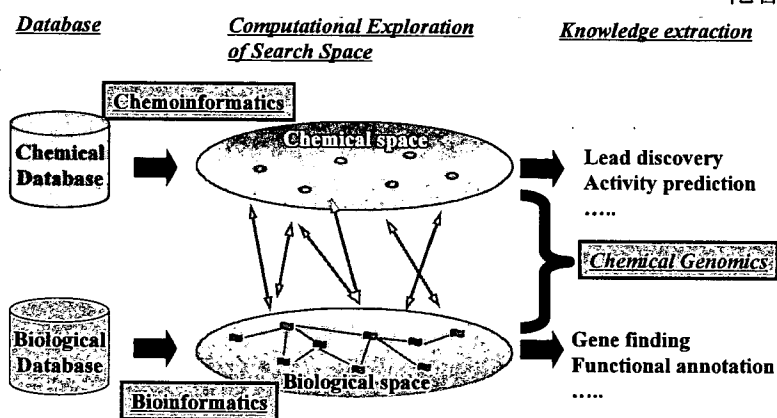
3. ケモインフォマティクスとバイオインフォマティクス(図1)

ケミカルゲノミクス・ケミカルバイオロジーでは、化合物のケミカル情報と生体系のバイオ情報の2種の異なる情報が対象となる。従って、ケミカルゲノミクスのための情報処理技術には、ケミカル情報を処理するケモインフォマティクスとバイオ情報を処理するバイオインフォマティクスを融合する新たなインフォマティクス技術の開発が必須となる。しかしながら、化学と生物学という異なる分野を背景にもつ2つのインフォマティクスは、独立して発展してきており現状では互いに相容れない。そこで、我々はケモインフォマティクスとバイオインフォマティクスにおける方法論的なアナロジーに着目しその融合を図った。すなわち、ケモインフォマティクスもバイオインフォマティクスも共に、個体(化合物やタンパク質)の特徴量を数値やベクトルで表現することにより、各個体の相対的な特性の違いを探索空間上の個体間の距離として定量的に算出する方法論を基本としている。例えば、ケモインフォマティクスでは、データベースに集積された膨大な化合物エントリーは化学構造や特性を定量的に表すベクトルとして表現され、その相対的な違いを距離の尺度として持つ座標空間(探索空間)をコンピュータ内部に構築する。データベース検索はこの探索空間において距離が近接する化合物を類似化合物として選出してくることに

なる。また、バイオインフォマティクスでも同様の考え方であり、遺伝子・タンパク質エントリーは配列や構造として表現され、それぞれの相同性（類似度）を尺度として持つ探索空間（バイオデータの場合、探索空間は系統樹やネットワーク構造になっている場合もある。）が構築され、データベース検索にはこの探索空間に基づき、類似（類縁）遺伝子・タンパク質が選出される。

一方、ケミカルゲノミクスとは、ケミカル空間の個体（化合物）とバイオ空間の個体（遺伝子・タンパク質）との相互作用関係を網羅的に明らかにする研究であり、図1の赤線に示す対応関係を付加したモデルであると考えられる。ここで、我々は、ケミカル情報とバイオ情報を統合的に処理するために、ケミカル空間（緑色）とバイオ空間（黄色）を独立して扱うのではなく、2つの空間を融合したモデルをケミカルゲノミクスのためのインフォマティクスモデルとして初めて考案した。

図1. ケモインフォマティクスとバイオインフォマティクス



4. ケミカル空間-バイオ空間の融合モデル

情報科学的アプローチによる化合物探索は、これまで化合物のケミカル情報のみを用いたケモインフォマティクス手法が用いられてきた。これに対し、我々の手法は、このケミカル情報のみの従来手法にバイオインフォマティクス技術を融合させ、バイオ情報を考慮に入れた化合物探索を実現する新しいインフォマティクス手法と言える。（図1）

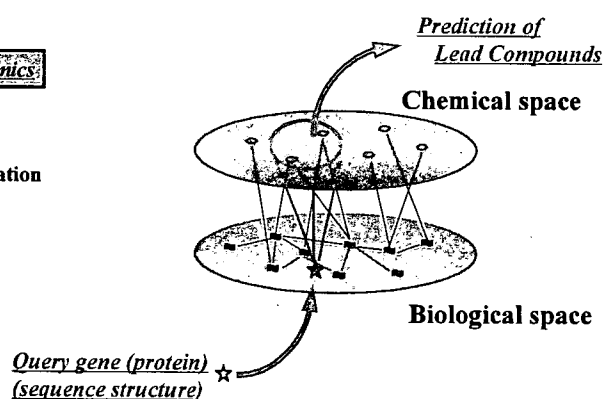
例えば、化合物について構造や特性の類似性を相対的な位置関係として表現したものをケミカル空間（赤が化合物、緑領域がケミカル空間）と

して定義するとともに、タンパク質についても類似関係（配列や構造の相同性）を相対的な位置関係として表現したものをバイオ空間（青がタンパク質、黄色領域がバイオ空間）として定義する。さらに個々の化合物とタンパク質の結合をリンク（黒線）することによって、これらケミカル空間とバイオ空間を融合した単純モデルをコンピュータ内部に構築できる。（図2）

ここで、標的タンパク質に作用する化合物候補を探索する In silico スクリーニングにこの融合モデルを適用する場合を考えると、

- 1) 標的タンパク質（青星）の配列構造から、そのタンパク質がバイオ空間座標にマッピングされる。
- 2) バイオ空間にマッピングされた標的タンパク質の近隣タンパク質からのケミカル空間へのリンク情報をたどること（青矢印）により、その標的タンパク質が関係するケミカル空間のエリア（青円内）を指定することができる。
- 3) 上記エリア内の化合物群が、標的タンパク質に相互作用する可能性のある化合物群と推定される。（ここでは、類似のタンパク質は、類似の化合物を結合するという前提を基にしている。）

図2. ケミカル空間-バイオ空間融合モデルを用いた In silico スクリーニング



我々は、このケミカル空間とバイオ空間の融合モデルを用いた探索を、GPCR ファミリーとそのリガンド化合物の探索に適用し、GLIDA データベース(2,3)として Web サービスを行っている。

5. ケミカル空間-バイオ空間の相関モデル

我々が考案する融合モデルで最も重要なことは、モデルの構築方法である。我々の方法は、こ

のモデル構築方法においても、従来法より大きな優位性を有している。

ケミカル空間の構築に用いられる従来法の代表的なものに主成分分析 (PCA) がある。これは、ケミカル情報のみを用い、化合物の化学特性ができる限り多様になるように、ケミカル空間座標を定義するものであり、ここでの大きな問題点は、化合物の多様性と生物活性との直接の因果関係は無いということである。

これらの問題点を克服することを目的とし、我々は正準相関分析 (CCA) を用いて相関モデル構築を試みた。本手法は、ケミカル情報とタンパク配列情報の両情報を用いて、ケミカル空間とタンパク空間の両空間の相関が高くなるように互いの空間座標を定義するものである。これはバイオ空間の分布を考慮して、ケミカル空間座標を定義する方が、生物活性にとって都合の良い空間座標を構築できるという大きな特徴を有する。

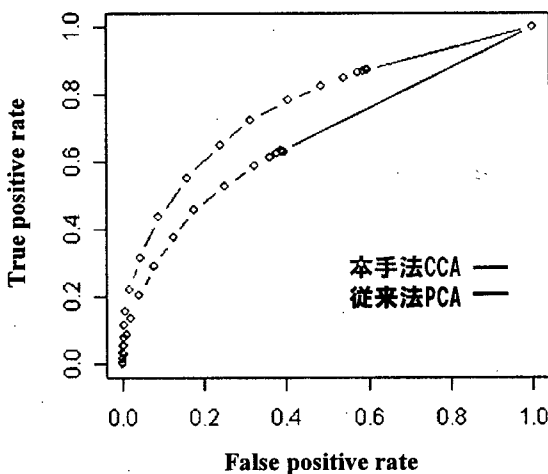


図3. 本手法と従来法との性能比較

実際に、従来法(PCA)と本手法(CCA)の性能を比較するために、既知データ(DrugBank データベース(4))を用いた 5-fold cross validation テストを実施し、化合物とタンパク質の相互作用予測の予測性能評価を行った。その結果を図3に示す。図は、予測性能を評価する有名な方法の一つであるROC 曲線 (横軸は化合物-タンパク質相互作用を誤って予測した割合、縦軸は正しく予測した割合を示す。) であり、このグラフは曲線が上に位置するほど、予測性能が良いことを表す。本手法の赤曲線が従来法の黒曲線より上方に位置するこ

とから、本手法の方が、従来法よりも高い予測性能であることがわかる。

6. Biologically relevant chemical space

ごく最近、宇宙観測において、暗黒物質 (ダークマター: 目には見えない物質) の宇宙空間での分布が初めて観測され宇宙の起源解明に期待が寄せられている。化合物の宇宙空間では、我々は何を探索しなければならないのであろうか? Chemical Space 特集号では、その答えとして「生物にとって意味のある化合物群 (Biologically relevant chemical space)」を探索することの重要性を提示している。しかしながら、Biologically relevant chemical space を定義する具体的方法に関する報告は未だなされていない。

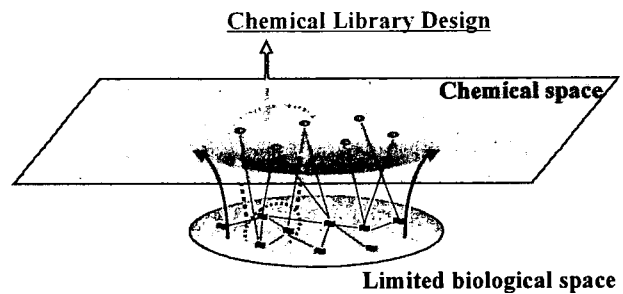


図4. Biologically relevant chemical space

我々の開発した融合モデルはこの難問に一つの解答を与える事ができ、タンパク質群と化合物群の相互作用様式の統計モデルを構築することにより、広大なケミカル空間のうち黄色のバイオ空間に対応するケミカル空間 (biologically relevant chemical space) を限定することを可能にする。(図4) また、バイオ空間のタンパク群を、例えば GPCR ファミリー (赤点線エリア) などに限定することにより GPCR 用のフォーカスライブラリー的设计も可能になり、生物活性を有する化合物ライブラリーの合理的設計が実現できる。

例えば、図5は、カナダでサービスされている DrugBank データベース(4)における 3476 個の既知薬物ターゲットタンパク質からなるバイオ空間とそれに対応する 3079 個の既知薬物からなるケミカル空間の実際の相関モデルである。このように、Biologically relevant chemical space の定義が可能となる。

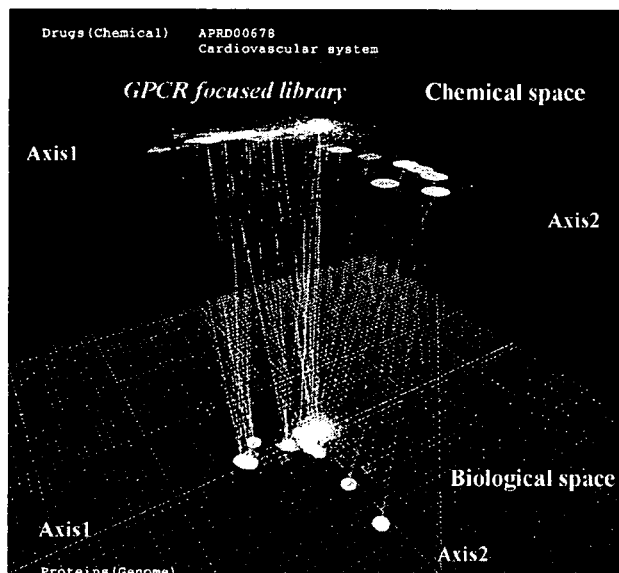


図5. DrugBank データを用いた実際のケミカル空間とバイオ空間

7. おわりに

医薬品の開発プロセスにおいて、現在用いられている In silico 技術には、我々が研究開発を行う情報科学的アプローチの他に、立体構造モデルを用いたドッキングシミュレーションのような計算化学的アプローチが有名である。情報科学的アプローチと計算化学的アプローチには、それぞれ一長一短があるが、欠点を互いに補完し合い In silico 創薬の確度向上を図ることが今後の課題であろう。特に、ケミカルゲノミクスが盛んな今日、日々増加し続ける莫大なデータを処理することは必須であり、情報科学的アプローチである創薬インフォマティクスのさらなる研究開発が必要である。

謝辞

本研究の一部は、文部科学省、厚生労働省の支援によって行われている。また、検証実験等の共同研究を行って頂いている京都大学薬学研究科ゲノム創薬科学分野の辻本豪三教授に深く感謝申し上げます。

参考文献

- 1) Nature, 432 (7019) (Insight), 823-865
- 2) Okuno, Y., Yang, J., Taneishi, K., Yabuuchi, H., Tsujimoto, G., Nucleic Acids Research, 34, D673-677
- 3) <http://pharminfo.pharm.kyoto-u.ac.jp/services/glida/>
- 4) <http://redpoll.pharmacy.ualberta.ca/drugbank/>

◆略歴◆ 奥野 恭史 (Yasushi OKUNO) : 1995年京大薬・修士課程修了、京大薬・博士後期課程進学、1996年京大薬・博士後期課程中途退学、京都大学化学研究所教務職員、2000年京大薬博士号取得、2001年京都大学化学研究所博士研究員、2002年京都大学化学研究所特任助手、2003年京都大学大学院薬学研究科特任助手、2005年産業技術総合研究所外来研究員(併任)、2006年京都大学大学院薬学研究科助教授

奨励賞受賞研究

バイオ空間とケミカル空間の包括的相関解析とそのインシリコ創薬への研究展開

京都大学大学院薬学研究科准教授 奥野 恭史



ポストゲノム時代の今日、莫大なゲノム情報から創薬への手がかりを発見すること、すなわち「ゲノム創薬」に大きな期待が寄せられてきた。ごく最近では、ゲノムプロジェクトで集積された大量のバイオ情報を低分子化合物へと直結することを目標とした「ケミカルゲノミクス」や「ケミカルバイオロジー」が注目さ

れ、生物学分野に加え化学分野にまで包括的概念が導入されつつある。実際、米国の筆頭に国内外において、ケミカルゲノミクスプロジェクトが開始され、「莫大な数の化合物と生体系との相互作用を包括的に明らかにすること」を命題に数百万もの膨大な化合物に関する情報の収集と有用化合物の探索が展開されている。

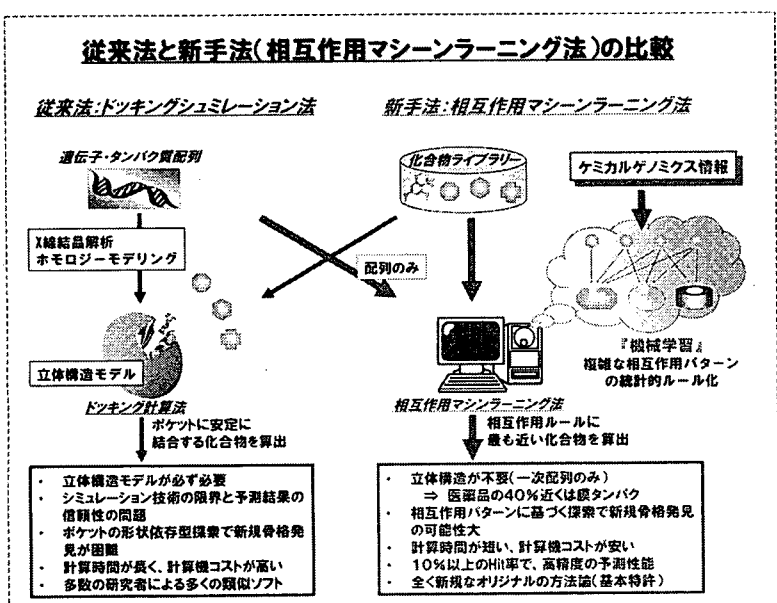
しかしながら、プロジェクトで算出される成果物は膨大なデータ量であるだけに、それらから生物学と化学とのインターフェイスとなる物理化学的法則性を発見しなければ、ケミカルゲノミクスが単なるやみくもなデータ収集に終わりがかねないという危険性を有している。

このような背景を鑑み、われわれは、莫大な化合物群とタンパク質群との相互作用様式をゲノムスケールで解析することに主眼を置いた情報科学に基づき理論的アプローチを世界に先駆けて研究開発し、ケミカルゲノミクスデータからの分子認識における物理化学的法則の発見と、その知識

に基づき新たなインシリコ創薬手法の開発を行ってきた。

代表的な研究成果としては、ケミカル情報とバイオ情報を統合的に処理するために、ケミカル空間とバイオ空間を独立して扱うのではなく、二つの空間を化合物とタンパク質との網羅的な相互作用関係に基づいて融合したモデルを、ケミカルゲノミクスのためのインフォマティクスモデルとして初めて考案した。また、この融合モデルのインシリコ創薬応用として、ケミカルゲノミクス情報の

に基づく新たなインシリコ創薬手法の開発を行ってきた。



機械学習に基づくインシリコスクリーニング手法の開発を行った。

この方法論は、従来のタンパク質の立体構造モデルを用いたドッキング

展させ、定量的な導入、構造活性相関情報の抽出などに取り組んでいる。

計算とは異なり、立体構造情報の代わりケミカルゲノミクス情報を用いて予測を行うという根本的に発想の異なるオリジナルな方法論であり、本手法をGPCRのリガンド探索に適用した結果、優れたヒット率と新規骨格探索能力を有することが証明された。現在は、ケミカル空間・バイオ空間融合モデルをさらに発

YAKU学・最前線

研究現場からあなたへ

ヒトゲノムプロジェクトが開始された当初、ゲノム解説が多くの新薬を生み出し、多くの疾患が治療可能になるものと、大きな夢と希望が抱かれました。しかし、残念ながら、ゲノム解説が宣言がなされた現在も、夢の医療の実現には至っていません。

とその所在ゲノム上の位置)が明らかになったことが挙げられます。すなわち、研究や医療の対象となり得る遺伝子が有限個に限定されたことにより、病気の原因や薬の治療標的となる遺伝子・蛋白質を徹底的に探し回る必要性がなくなったのです。生命科学の研究を展開する際に、つぎの2000ほどの遺伝子(下の全遺伝子)及びそれらの組み合わせを対象にすればよく、目指すべき対象とその範囲が明確になったというメリットは計り知れないものだと思います。

またさらに、ゲノムプロジェクトは、科学の世界に、生命現象の全体像を捉えることの重要性、総体、包括性、網羅性と

す。ごく最近、この包括的価値観は、化学(ケミストリー)の世界にも導入され始めています。遺伝

数はいったい幾つあるのでしょうか? 進化の過程で淘汰されてきた遺伝子とは異なり、化合物は無

に存在する無限の星に比べて、化合物の宇宙(ケミカル宇宙、ケミカルスペース)と称するのは、

このことは、広大な宇宙の中から未知の惑星を発見する宇宙探索と同様に、無限のケミカル宇宙から医薬品の種となる活性化化合物を

発見することが非常に困難なことを物語っています。近年、これらケミカル空間を探索する研究分野として、ケミカルゲノミクスやケミカルバイオロジイという分野が注目されています。ケミカルゲノミクスでは、その命題として「莫大な数の化合物と生体系(蛋白質や細

菌など)との相互作用を包括的に明らかにすることが掲げられ、米国では、数百万もの膨大な化合物に関する情報を収集し、有用な化合物を探索する国家プロジェクトが展開されています。しかしながら、有限個数を対象にしていたゲノムプロジェクトと異なり、ケミカルゲノミクスプロジェクトでいかに網羅的研究展開がなされたとしても、無限のケミカル空間を探索し尽くすことは不可能であります。われわれの研究室(Chris / Pharming / Pharma Probe, ac. Inc.)では、この無限に広がる化合物の宇宙空間をコンピュータ上にモデル化し、情報科学技術を用いて医薬品の種となる活性化化合物の探索を行っています。コンピュータの中では、多次元空間を構築することが可能であり、無限の化合物の可能性を対象にできるからです。化合物が無数であるということは、未知の化合物、すばらしい薬効を有する化合物が存在している可能性も無限であります。

ゲノムからケミカル宇宙へ

では、ゲノムプロジェクトは無意味なものだったのでしょうか?

もちろん、答えはNoです。例えば、ゲノムプロジェクトの大きな貢献の一つに、ヒトを形作る全ての

この新しい価値観を根づかせました。すなわち、個々の遺伝子、蛋白質、分子の働きを見るのが主流とされてきたこれまで

子の総体をゲノムと名づけるのに対し、化合物の総体はケミカル空間(ケミカル宇宙、ケミカルスペース)と総称されています。

では、ゲノムプロジェクトと同様に、ケミカル空間を構成する化合物数、すなわち化合物の総

数の組み合わせを創製することが可能であることから、無限に存在すると考えられます。例えば、分子数500以下のものだけども、理論上では二億種類を超える天文学的な

推定されています。従って、無数の化合物を宇宙

宇宙の中に、地球と同じ惑星を探すと同等様に、ケミカル宇宙の中の新規化合物探索は夢とロマンに満ち溢れています。

もちろん、答えはNoです。例えば、ゲノムプロジェクトの大きな貢献の一つに、ヒトを形作る全ての

この新しい価値観を根づかせました。すなわち、個々の遺伝子、蛋白質、分子の働きを見るのが主流とされてきたこれまで

子の総体をゲノムと名づけるのに対し、化合物の総体はケミカル空間(ケミカル宇宙、ケミカルスペース)と総称されています。

では、ゲノムプロジェクトと同様に、ケミカル空間を構成する化合物数、すなわち化合物の総

数の組み合わせを創製することが可能であることから、無限に存在すると考えられます。例えば、分子数500以下のものだけども、理論上では二億種類を超える天文学的な

推定されています。従って、無数の化合物を宇宙

宇宙の中に、地球と同じ惑星を探すと同等様に、ケミカル宇宙の中の新規化合物探索は夢とロマンに満ち溢れています。

宇野 啓史
京都大学薬学研究所
統合薬学フロンティア
ア教育センター准教授