Table 7 Analysis of divergence of the Shi470-type *dupA*

	Nucleotide and amino acid identity[a] (%)													
Strain	Shi470	F51	F57	F58	F64	F77	F228	OK99	OK108[b]	OK165	OK169	OK203	OK303	OK309
Shi470		98.2	98.8	99.0	99.0	98.3	98.9	98.3	–	98.1	98.1	98.4	98.8	98.2
F51	98.3		98.2	98.2	98.2	98.0	98.3	97.7	–	97.7	97.6	97.8	98.2	97.6
F57	98.6	98.4		99.0	99.0	98.8	99.4	98.8	–	98.8	98.8	99.2	99.3	98.9
F58	98.8	98.4	98.7		100	98.6	99.2	98.8	–	98.6	98.3	98.7	99.0	98.4
F64	98.8	98.4	98.7	100		98.6	99.2	98.8	–	98.6	98.3	98.7	99.0	98.4
F77	98.7	98.6	99.0	98.8	98.8		98.7	99.0	–	98.3	98.1	98.4	98.6	98.2
F228	98.8	98.4	99.4	98.8	98.8	98.9		98.7	–	98.7	98.7	99.0	99.6	98.8
OK99	98.5	98.2	98.8	98.9	98.9	99.1	98.8		–	98.6	98.1	98.4	98.4	98.2
OK108[b]	98.5	98.2	99.2	99.1	99.1	98.9	99.2	98.4		–	–	–	–	–
OK165	98.5	98.5	98.9	98.9	98.9	99.0	98.9	99.3	98.6		99.3	99.4	98.6	99.4
OK169	98.2	98.2	99.0	98.3	98.3	98.6	99.1	98.6	98.7	99.2		99.6	98.6	99.9
OK203	98.3	98.3	99.1	98.5	98.5	98.8	99.2	98.7	98.9	99.2	99.8		98.9	99.8
OK303	98.7	98.4	99.4	98.8	98.8	98.9	98.8	98.8	99.2	98.9	99.1	99.2		98.7
OK309	98.2	98.2	99.0	98.4	98.4	98.7	99.1	98.6	98.8	99.2	100	99.8	99.1	

dupA, duodenal ulcer-promoting gene A; a values below and above the diagonal represent nucleotide and amino acid sequence, respectively; b a stop codon is present in the *dupA* nucleotide sequence.

dupA prevalence in *H. pylori* strains from Belgium, South Africa, China, and the United States. These investigators reported a significant association between *dupA* positivity with GC, independent of *cagA* status.[32] We previously showed that more than 20% of Okinawan strains were either Western-type *cagA*$^+$ or *cagA*$^-$, both of which were rare in East Asia.[23] We also found these Western-type strains were isolated among patients with DU more frequently than the East Asian strains,[26,27] whereas the East Asian-type strains in Okinawa were associated with GC.[27] Therefore, we examined the association between *dupA* and *cagA* status in Okinawa. Intriguingly, we found a low prevalence of the *dupA*$^+$ strain among strains with the non-East Asian *cagA* (the Western or hybrid type) in Okinawa. Moreover, both univariate and multivariate analyses showed that the presence of East Asian-type *cagA* was significantly associated with *dupA* positivity. Although the reason why *dupA*$^+$ strains are rare among non-East Asian *cagA*$^+$ strains is unclear, the variety in *cagA* toxicity may partially, if not completely, explain this mystery. Hussein et al. reported that strains with *cagA* genes encoding more than 3 tyrosine phosphorylation motifs of the Src homology phosphatase 2 (SHP-2) binding site were significantly associated with *dupA* positivity when compared to strains with only 3 phosphorylation motifs in Iran.[35] The virulence of *cagA*, which is determined by the degree of CagA SHP-2 binding activity, depends on the genotypes of the SHP-2 binding site (the East Asian-type is more toxic than the Western-type) and the number of its phosphorylation motifs (the greater the number of motifs, the stronger the binding).[22,25,27,52] Together, the results of Hussein et al. and those of our present study suggest the possible association of *dupA* with "toxic" *cagA*.

We found that the *dupA* prevalence in Okinawa was significantly higher than in Fukui by using multivariable logistic regression. This result seems partially consistent with those of recent meta-analyses showing that the prevalence of *dupA* is higher in Western countries than in Asian countries,[50,51] as our previous reports suggested *H. pylori* strains in Okinawa may have had a greater opportunity for the transfer of DNA from Western countries than *H. pylori* strains in other Japanese areas.[23,26] Since this question has not been sufficiently studied, determining the genetic origin of *H. pylori* strains circulating in Okinawa is required in examining the roles of *H. pylori* virulent factors, including *dupA*, in this area.

In this study, we analyzed the upstream nucleotide sequences of *dupA* in Japanese clinical *H. pylori* isolates, which have not been sufficiently examined in Japan thus far. A comparison between the genome sequences of 12 representative *dupA*$^+$ strains of *H. pylori* in GenBank, including our strain (F57), revealed the existence of 2 major distinct chromosomal arrangements (Shi470 and J99 types). Regardless of the host's clinical outcome, we showed that most *dupA* sequences of the Japanese *H. pylori* strains were of the intact Shi470-type gene, which encodes a protein of 832 amino acids. Although this finding made it difficult to determine if *dupA* was a specific virulence factor in Japan, we demonstrated a strong association between intact *dupA* and *H. pylori* strains circulating in this country. When discussing the virulence of a particular *H. pylori* gene, it is important to consider not only the presence of the gene but also whether the gene is intact. We previously showed *vacA* from half of the Japanese non-cytotoxic strains contained null mutations resulting in a lack of VacA activity.[53] In terms of *dupA*, Queiroz et al. reported an inverse association between the presence of *dupA* without particular null mutations (a deletion of adenine at position 1311 and insertion of adenine at position 1426) and the development of GC in Brazil.[41]

Notably, Hussein et al. had classified *dupA* into 2 main groups (*dupA1* and *dupA2*) before us. This group reported *dupA1* is the most common genotype, as well as the active form that induces host secretion of IL-12 from cluster of differentiation 14 (CD14) positive mononuclear cells. Among the 25 *dupA1*$^+$ isolates studied by Hussein et al., additional sequencing was performed for 2 isolates. Surprisingly, these isolates possessed the *dupA* gene with an extended 5′ end, giving a total length of 2499 bp.[45] Because all 13 Shi470-type *dupA* strains sequenced

in this study included the 1884bp dupA1 defined by Hussein et al., *dupA1* may be longer than initially reported, corresponding to the Shi470-type *dupA* in the present study. On the other hand, we cannot deny the possibility the Shi470-type *dupA* is a third novel genotype of *dupA*, different from both *dupA1* and *dupA2*. Following Hussein et al., Queiroz et al. sequenced the complete *dupA* gene from 6 *H. pylori* isolates in Brazil. Though all these genes were reported to be identical to *dupA1*, the report did not specify whether the 5′ end was extensive or not.[46] For assessing the correspondence of the Shi470-type *dupA* to *dupA1*, further investigations on the upstream region of *dupA* are required. Investigating the *dupA* genotype might provide insight into the potential relationship between intact *dupA* and the strong virulence of *H. pylori* strains.

It should be noted that *dupA* gene expression has not been sufficiently analyzed thus far. Only 2 of the previous studies on *dupA* have investigated its expression in various isolates, as determined by reverse transcription PCR (RT-PCR). In Japan, Nguyen et al. confirmed *dupA* expression in all 10 randomly selected *H. pylori* strains among 72 *dupA*$^+$ isolates.[39] Only recently, Alam et al. reported a significant association of *dupA* with DU development in a South Indian population.[43] They also reported the detection of *dupA* transcription in 28 of 35 *dupA*$^+$ *H. pylori* strains. The discovery of strains with "expressive *dupA*" by RT-PCR and an assessment of their relationship with clinical outcomes would reveal the true role of the DupA protein in the disease development of the host. Therefore, we should next assess the *dupA* expression of the Japanese isolates used in the present study, especially the Shi470-type *dupA*$^+$ isolates, by using the methods of Alam et al. Furthermore, the detection of DupA protein still remains an unsolved problem.[30]

Further exploration of the surrounding genes of *dupA* may also be promising for understanding *H. pylori* virulence. Previous studies have shown that *dupA* is present in the plasticity zone. This zone is characterized by a low G+C content and an abundance of strain-specific ORFs, including TFSSs.[11,28,29] Recently, Kersulyte et al. described a putative T4SS-containing *dupA* as a type IV secretion 3a (*tfs3a*) gene, and a putative T4SS-containing a *virB4* sequence, but not *dupA* as *tfs3b*.[29] Jung et al. defined the presence of *dupA* and all 6 adjacent virulence (*vir*) gene homologs (*virB8*, *B9*, *B10*, *B11*, *virD4*, and *D2*) as a complete *dupA* cluster.[30] They reported that the presence of a complete *dupA* cluster but not *dupA* alone is associated with DU development in the United States population. Because we could not find a particular association between *dupA* and any clinical outcomes, the *vir* genes around *dupA* in the Japanese *H. pylori* isolates studied in the present study will need to be researched further.

In conclusion, we were unable to confirm a close association between *dupA* and clinical outcomes in 2 areas in Japan with different GC risks. However, we found intact *dupA* consists of a 2499 bp sequence and most Japanese *dupA*$^+$ strains possess the intact genotype. We also discovered a significant relation between the presence of *dupA* and the East Asian-type *cagA* in Japan. In addition, *dupA* positivity was independently associated with strains isolated from Okinawa.

Further global analyses of the genetic structure of *dupA* and the association between the *H. pylori* strain with a *dupA*-positive state and clinical outcomes are required.

Contributors

Hidetaka Matsuda, Yoshiyuki Ito, Hiroyuki Suto, Satoko Satomi, Masahiro Ohtani, Yukinao Yamazaki, Yoshiki Shimabukuro, Kaoru Kikuchi, Yoshihide Keida, Takeshi Azum, and Yasunari Nakamoto were involved in conceiving and designing this study. Hidetaka Matsuda, Yoshiyuki Ito, and Akiyo Yamakawa analyzed and interpreted the data. Yukinori Kusaka supervised and revised the statistics. This manuscript was drafted by Hidetaka Matsuda, revised critically by Yoshiyuki Ito for important intellectual content, and finalized by Yasunari Nakamoto with the assistance of all authors.

Conflicts of interest

The authors declared no conflicts of interest.

Summary Box

What is already known:

- The association between duodenal ulcer-promoting gene A (*dupA*) of *Helicobacter pylori* (*H. pylori*) and various clinical outcomes or other virulence factors remains controversial.
- In Japan, Okinawa has lowest incidence of gastric cancer (GC). The diversity of cytotoxin-associated gene A (*cagA*) detected in Okinawa includes both East Asian and Western genotypes and is an important factor for the low prevalence of GC.
- *dupA* is reported to be polymorphic; *dupA1* (1884 bp) is the active form, while *dupA2* is the truncated version. Several *H. pylori* strains are reported to possess long *dupA1* sequences (2499 bp) because of an extended 5′ end.

What the new findings are:

- We found *dupA* could be classified into 2 genotypes—a 2499 bp genotype and the initially reported 1839 bp genotype. These genotypes were designated Shi470-type *dupA* and J99-type *dupA*, respectively, after the *H. pylori* strains whose complete genomes have been released.
- In Japan, most *dupA*$^+$ *H. pylori* strains possess intact Shi470-type *dupA*.
- In Japan, *dupA* appeared to be independently associated with the East Asian-type *cagA* and with host residence in Okinawa.

References

1. Marshall BJ, Goodwin CS, Warren JR, et al. Prospective double blind-trial of duodenal ulcer relapse after eradication of *Campylobacter pylori*. Lancet 1988; 2(8626):1437-42.
2. Parsonnet J, Friedman GD, Vandersteen DP, et al. *Helicobacter pylori* infection and the risk of gastric carcinoma. N Engl J Med 1991; 325(16):1127-31.
3. Wotherspoon AC, Ortiz-Hidalgo C, Falzon MR, et al. *Helicobacter pylori*-associated gastritis and primary B-cell gastric lymphoma. Lancet 1991; 338(8776):1175-6.
4. Blaser MJ, Perez-Perez GI, Kleanthous H, et al. Infection with *Helicobacter pylori* strains possessing cagA is associated with an increased risk of developing adenocarcinoma of the stomach. Cancer Res 1995; 55(10):2111-5.
5. Parsonnet J, Friedman GD, Orentreich N, et al. Risk for gastric cancer in people with CagA positive or CagA negative Helicobacter pylori infection. Gut 1997; 40(3):297-301.
6. Stephens JC, Stewart JA, Folwell AM, et al. *Helicobacter pylori cagA* status, *vacA* genotypes and ulcer disease. Eur J Gastroenterol Hepatol 1998; 10(5):381-4.
7. Hamlet A, Thoreson AC, Nilsson O, et al. Duodenal *Helicobacter pylori* infection differs in *cagA* genotype between asymptomatic subjects and patients with duodenal ulcers. Gastroenterology 1999; 116(2):259-68.
8. Rugge M, Busatto G, Cassaro M, et al. Patients younger than 40 years with gastric carcinoma: *Helicobacter pylori* genotype and associated gastritis phenotype. Cancer 1999; 85(12):2506-11.
9. Atherton JC, Cao P, Peek RM Jr, et al. Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific *vacA* types with cytotoxin production and peptic ulceration. J Biol Chem 1995; 270(30):17771-7.
10. Atherton JC, Peek RM Jr, Tham KT, et al. Clinical and pathological importance of heterogeneity in *vacA*, the vacuolating cytotoxin gene of *Helicobacter pylori*. Gastroenterology 1997; 112(1):92-9.
11. Yamaoka Y. Roles of the plasticity regions of *Helicobacter pylori* in gastroduodenal pathogenesis. J Med Microbiol 2008; 57(5):545-53.
12. Wen S, Moss SF. *Helicobacter pylori* virulence factors in gastric carcinogenesis. Cancer Lett 2009; 282(1):1-8.
13. Proença-Modena JL, Acrani GO, Brocchi M. *Helicobacter pylori*: phenotypes, genotypes and virulence genes. Future Microbiol 2009; 4(2):223-40.
14. Covacci A, Telford JL, Del Giudice G, et al. *Helicobacter pylori* virulence and genetic geography. Science 1999; 284(5418):1328-33.
15. Falush D, Wirth T, Linz B, et al. Traces of human migrations in *Helicobacter pylori* populations. Science 2003; 299(5612):1582-5.
16. Linz B, Balloux F, Moodley Y, et al. An African origin for the intimate association between humans and *Helicobacter pylori*. Nature 2007; 445(7130):915-8.
17. Azuma T. *Helicobacter pylori* CagA protein variation associated with gastric cancer in Asia. J Gastroenterol 2004; 39(2):97-103.
18. Correa P. A human model of gastric carcinogenesis. Cancer Res 1988; 48(13):3554-60.
19. Ito S, Azuma T, Murakita H, et al. Profile of *Helicobacter pylori* cytotoxin derived from two areas of Japan with different prevalence of atrophic gastritis. Gut 1996; 39(6):800-6.
20. Nobuta A, Asaka M, Sugiyama T, et al. *Helicobacter pylori* infection in two areas in Japan with different risks for gastric cancer. Aliment Pharmacol Ther 2004; 20 (Suppl 1):1-6.
21. Ito Y, Azuma T, Ito S, et al. Sequence analysis and clinical significance of the *iceA* gene from *Helicobacter pylori* strains in Japan. J Clin Microbiol 2000; 38(2):483-8.
22. Azuma T, Yamakawa A, Yamazaki S, et al. Correlation between variation of the 3′ region of the *cagA* gene in *Helicobacter pylori* and disease outcome in Japan. J Infect Dis 2002; 186(11):1621-30.
23. Zhou W, Yamazaki S, Yamakawa A, et al. The diversity of *vacA* and *cagA* of *Helicobacter pylori* in East Asia. FEMS Immunol Med Microbiol 2004; 40(1):81-7.
24. Azuma T, Yamakawa A, Yamazaki S, et al. Distinct diversity of the cag pathogenicity island among *Helicobacter pylori* strains in Japan. J Clin Microbiol 2004; 42(6):2508-17.
25. Azuma T, Yamazaki S, Yamakawa A, et al. Association between diversity in the Src homology 2 domain-containing tyrosine phosphatase binding site of *Helicobacter pylori* CagA protein and gastric atrophy and cancer. J Infect Dis 2004; 189(5):820-7.
26. Yamazaki S, Yamakawa A, Okuda T, et al. Distinct diversity of *vacA*, *cagA*, and *cagE* genes of *Helicobacter pylori* associated with peptic ulcer in Japan. J Clin Microbiol 2005; 43(8):3906-16.
27. Satomi S, Yamakawa A, Matsunaga S, et al. Relationship between the diversity of *cagA* gene of *Helicobacter pylori* and gastric cancer in Okinawa, Japan. J Gastroenterol 2006; 41(7):668-73.
28. Lu H, Hsu PI, Graham DY, et al. Duodenal ulcer promoting gene of *Helicobacter pylori*. Gastroenterology 2005; 128(4):833-48.
29. Kersulyte D, Lee W, Subramaniam D, et al. *Helicobacter pylori's* plasticity zones are novel transposable elements. PLoS One 2009; 4(9):e6859.
30. Jung SW, Sugimoto M, Shiota S, et al. The intact *dupA* cluster is a more reliable *Helicobacter pylori* virulence marker than *dupA* alone. Infect Immun 2012; 80(1):381-7.
31. Arachchi HS, Kalra V, Lal B, et al. Prevalence of duodenal ulcer-promoting gene (*dupA*) of *Helicobacter pylori* in patients with duodenal ulcer in North Indian population. Helicobacter 2007; 12(6):591-7.
32. Argent RH, Burette A, Miendje Deyi VY, et al. The presence of *dupA* in *Helicobacter pylori* is not significantly associated with duodenal ulceration in Belgium, South Africa, China, or North America. Clin Infect Dis 2007; 45(9):1204-6.
33. Gomes LI, Rocha GA, Rocha AM, et al. Lack of association between *Helicobacter pylori* infection with *dupA*-positive strains and gastroduodenal diseases in Brazilian patients. Int J Med Microbiol 2008; 298(3-4):223-30.
34. Douraghi M, Mohammadi M, Oghalaie A, et al. *dupA* as a risk determinant in *Helicobacter pylori* infection. J Med Microbiol 2008; 57(5):554-62.
35. Hussein NR, Mohammadi M, Talebkhan Y, et al. Differences in virulence markers between *Helicobacter pylori* strains from Iraq and those from Iran: potential importance of regional differences in *H. pylori*–associated disease. J Clin Microbiol 2008; 46(5):1774-9.
36. Pacheco AR, Proença-Módena JL, Sales AI, et al. Involvement of the *Helicobacter pylori* plasticity region and cag pathogenicity island genes in the development of gastroduodenal diseases. Eur J Clin Microbiol Infect Dis 2008; 27(11):1053-9.
37. Zhang Z, Zheng Q, Chen X, et al. The *Helicobacter pylori* duodenal ulcer promoting gene, *dupA* in China. BMC Gastroenterol 2008; 8:49.
38. Schmidt HM, Andres S, Kaakoush NO, et al. The prevalence of the duodenal ulcer promoting gene (*dupA*) in *Helicobacter pylori* isolates varies by ethnic group and is not universally associated with disease development: a case-control study. Gut Pathog 2009; 1(1):5.
39. Nguyen LT, Uchida T, Tsukamoto Y, et al. *Helicobacter pylori dupA* gene is not associated with clinical outcomes in the Japanese population. Clin Microbiol Infect 2010; 16(8):1264-9.
40. Imagawa S, Ito M, Yoshihara M, et al. *Helicobacter pylori dupA* and gastric acid secretion are negatively associated with gastric cancer development. J Med Microbiol 2010; 59(12):1484-9.
41. Queiroz DM, Rocha GA, Rocha AM, et al. *dupA* polymorphisms and risk of *Helicobacter pylori*-associated diseases. Int J Med Microbiol 2011; 301(3):225-8.
42. Abadi AT, Taghvaei T, Wolfram L, et al. Infection with *Helicobacter pylori* strains lacking *dupA* is associated with an increased risk of gastric ulcer and gastric cancer development. J Med Microbiol 2012; 61(1):23-30.
43. Alam J, Maiti S, Ghosh P, et al. Significant association of *dupA* of *Helicobacter pylori* with duodenal ulcer development in South East Indian population. J Med Microbiol 2012; 61(Pt 9):1295-302.
44. Matteo MJ, Armitano RI, Granados G, et al. *Helicobacter pylori oipA*, *vacA* and *dupA* genetic diversity in individual hosts. J Med Microbiol 2010; 59(1):89-95.
45. Hussein NR, Argent RH, Marx CK, et al. *Helicobacter pylori* dupA is polymorphic, and its active form induces proinflammatory cytokine secretion by mononuclear cells. J Infect Dis 2010; 202(2):261-9.
46. Queiroz DM, Moura SB, Rocha AM, et al. The genotype of the Brazilian *dupA*-positive *Helicobacter pylori* strains is *dupA1*. J Infect Dis 2011; 203(7):1033-4.
47. Ito Y, Azuma T, Ito S, et al. Analysis and typing of the *vacA* gene from *cagA*-positive strains of *Helicobacter pylori* isolated in Japan. J Clin Microbiol 1997; 35(7):1710-4.
48. Mizushima T, Sugiyama T, Komatsu Y, et al. Clinical relevance of the *babA2* genotype of *Helicobacter pylori* in Japanese clinical isolates. J Clin Microbiol 2001; 39(7):2463-5.
49. Avasthi TS, Devi SH, Taylor TD, et al. Genomes of two chronological isolates (*Helicobacter pylori* 2017 and 2018) of the West African *Helicobacter pylori* strain 908 obtained from a single patient. J Bacteriol 2011; 193(13):3385-6.
50. Hussein NR. The association of *dupA* and *Helicobacter pylori*-related gastroduodenal diseases. Eur J Clin Microbiol Infect Dis 2010; 29(7):817-21.
51. Shiota S, Matsunari O, Watada M, et al. Systematic review and meta-analysis: the relationship between the *Helicobacter pylori dupA* gene and clinical outcomes. Gut Pathog 2010; 2(1):13.
52. Higashi H, Tsutsumi R, Fujita A, et al. Biological activity of the *Helicobacter pylori* virulence factor CagA is determined by variation in the tyrosine phosphorylation sites. Proc Natl Acad Sci USA 2002; 99(22):14428-33.
53. Ito Y, Azuma T, Ito S, et al. Full-length sequence analysis of the *vacA* gene from cytotoxic and noncytotoxic *Helicobacter pylori*. J Infect Dis 1998; 178(5):1391-8.

Cancer Gene Therapy (2012) **19,** 312–319

www.nature.com/cgt

ORIGINAL ARTICLE

Membrane-bound form of monocyte chemoattractant protein-1 enhances antitumor effects of suicide gene therapy in a model of hepatocellular carcinoma

Y Marukawa[1], Y Nakamoto[1], K Kakinoki[1], T Tsuchiyama[1], N Iida[1], T Kagaya[1], Y Sakai[1], M Naito[2], N Mukaida[3] and S Kaneko[1]

Suicide gene therapy using the herpes simplex virus thymidine kinase/ganciclovir (HSV-tk/GCV) system combined with monocyte chemoattractant protein-1 (MCP-1) provides significant antitumor efficacy. The current study was designed to evaluate the antitumor immunity of a newly developed membrane-bound form of MCP-1 (mMCP-1) in an immunocompetent mouse model of hepatocellular carcinoma (HCC). A recombinant adenovirus vector (rAd) harboring the human *MCP-1* gene and the membrane-spanning domain of the *CX3CL1* gene was used. Large amounts of MCP-1 protein were expressed and accumulated on the tumor cell surface. The growth of subcutaneous tumors was markedly suppressed when tumors were treated with mMCP-1, as compared with soluble MCP-1, in combination with the HSV-tk/GCV system ($P<0.01$). The numbers of Mac-1-, CD4- and CD8a-positive cells were significantly higher in tumor tissues ($P<0.05$), and tumor necrosis factor (TNF) mRNA expression levels with mMCP-1 were almost five-fold higher than those with soluble MCP-1. These results indicate that the delivery of the *mMCP-1* gene greatly enhanced antitumor effects following the apoptotic stimuli by promoting the recruitment and activation of macrophages and T lymphocytes, suggesting a novel strategy of immune-based gene therapy in the treatment of patients with HCC.

Cancer Gene Therapy (2012) **19,** 312–319; doi:10.1038/cgt.2012.3; published online 9 March 2012

Keywords: herpes simplex virus thymidine kinase; hepatocellular carcinoma; membrane-bound form; monocyte chemoattractant protein-1; monocyte/macrophages

INTRODUCTION

In spite of the recent development of locoregional treatments for hepatocellular carcinoma (HCC), the frequency of tumor recurrence remains high, probably because of insufficient therapeutic effects and the multicentric development of HCC in cirrhotic liver.[1–3] Non-surgical treatments of HCC, such as radiofrequency ablation, transcatheter arterial embolization and transcatheter arterial chemotherapy induce apoptosis of HCC cells, but these treatments do not enhance antitumoral immunity sufficiently. Thus, gene therapy aimed at enhancing antitumor immune responses may be a promising approach to prevent HCC recurrence, when it is combined with non-surgical maneuvers.

We previously reported that monocyte chemoattractant protein-1 (*MCP-1*) gene delivery using recombinant adenovirus vector (rAd) *in vivo* can enhance the efficacy of suicide gene therapy consisting of the delivery of rAd containing herpes simplex virus thymidine kinase (HSV-tk) and ganciclovir (GCV) in models of HCC[4,5] and colon cancer.[6] We further demonstrated that the antitumor effects depended on the activation of macrophages.[4,5] The adenovirus-specific spatial and temporal expression pattern may result in the production of the transgene for a limited time.[7] Mirroring these characteristics, the adenovirus vector-mediated delivery of the *MCP-1* gene alone was not sufficient to reduce tumor growth.[5] To circumvent this bottleneck, sustained expression of MCP-1 at the tumor site may be required to enhance the efficacy of the gene therapy using the *MCP-1* gene.

Systemic or local administration of cytokines has been used to enhance the antitumor immune response induced by many cancer vaccines. However, the systemic administration of cytokines resulted in unwanted side effects. Recently, tumor therapy that uses a membrane-bound form of cytokine was developed to reduce the side effects of cytokine in the systemic circulation. These experiments revealed that the membrane-bound form of cytokine not only reduced the side -effects, but enhanced the antitumor effects by prolonging the half-life of cytokines in the tumor microenvironment.[8]

These observations prompted us to design the adenovirus vector driving the expression of membrane-bound form of MCP-1 (mMCP-1) and to evaluate its antitumor effects in a model of HCC. We demonstrated that the delivery of the *mMCP-1* gene markedly augmented HSV-tk/GCV suicide gene therapy, compared with that of the soluble MCP-1 (sMCP-1).

MATERIALS AND METHODS

Recombinant adenovirus vectors

Ad-mMCP-1 (Figure 1a) harboring the human *MCP-1* gene and the membrane-spanning domain of the *CX3CL1* gene driven by the human cytomegalovirus

[1]Department of Disease Control and Homeostasis, Graduate School of Medical Science, Kanazawa University, Kanazawa, Japan; [2]Division of Cellular and Molecular Pathology, Niigata University Graduate School of Medicine, Niigata, Japan and [3]Division of Molecular Bioregulation, Cancer Research Institute, Kanazawa University, Kanazawa, Japan. Correspondence: Dr S Kaneko, Department of Disease Control and Homeostasis, Graduate School of Medical Science, Kanazawa University, 13-1 Takara-machi, Kanazawa 920-8641, Japan.
E-mail: skaneko@m-kanazawa.jp
Received 5 July 2011; revised 5 December 2011; accepted 26 January 2012; published online 9 March 2012

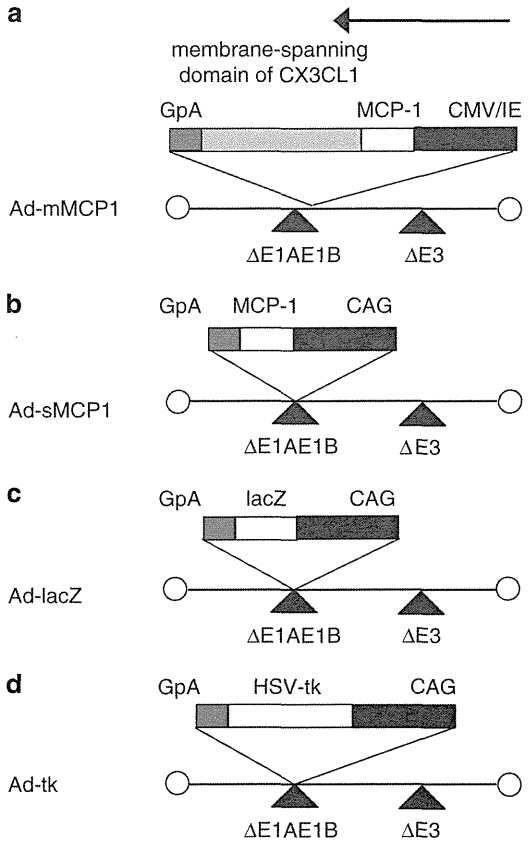

Figure 1. Construct of recombinant adenovirus vector (rAds). Under the control of the cytomegalovirus immediate early promoter/enhancer (CMV/IE) promoter and the CAG promoter, rAd Ad-membrane-bound monocyte chemoattractant protein-1 (mMCP-1) (**a**) expressing human MCP-1 and the membrane-spanning domain of fractalkine/CX3CL1 in sequence, rAd Ad-soluble-MCP-1 (sMCP-1), (**b**) expressing MCP-1, rAd Ad-lacZ, (**c**) expressing lacZ and rAd Ad-tk, (**d**) expressing HSV-tk. Solid lines indicate the rAd genome. An open triangle below the rAd genome represents a deletion of adenovirus early regions. Arrows show the orientations of the transcription. GpA, rabbit β-globin poly (A) site.

immediate early promoter/enhancer was prepared, purified and titrated according to the protocols supplied by the manufacturer (Takara, Tokyo, Japan). The human MCP-1/CX3CL1 (fractalkine) chimera was designed as follows: DNA encoding a fragment of human CX3CL1 spanning the intracellular, transmembrane and partial extracellular region was amplified from the full-length CX3CL1 cDNA by PCR with the following primers (5′-GCGAGCTCGGGTACCTTCGAGAAGCAGATCG-3′ and 5′-GCGAATTCAGATT GTCACACGGGCACAGG-3′). *Sac*I, *Kpn*I and *Eco*RI restriction enzyme sites were added at the 5′ and 3′ ends of this fragment. MCP-1 was also amplified by PCR with the following primers (5′-GCGAGCTCGCCAGCAT-GAAAGTCTCTGCCG-3′ and 5′-GCGGTACCAGTCTTCGGAGTTTGGGTTTGC-3′). *Sac*I and *Kpn*I restriction enzyme sites were added at the 5′ and 3′ ends of this fragment. The CX3CL1 and MCP-1 DNA fragments were digested with restriction enzymes by coligation into the *Sac*I and *Eco*RI sites of pSTBlue-1 (Novagen, Darmstadt, Germany), generating pSTBlue-1-mMCP-1. Then, pSTBlue-1-mMCP-1 was digested by *Not*I and *Bam*HI restriction enzymes and the fragment was inserted into the pShuttle Vector (Clontech Laboratories, Mountain View, CA) under the control of cytomegalovirus immediate early promoter/enhancer, generating pShuttle-mMCP-1. The pShuttle-mMCP-1 was digested with *PI-SceI/I-CeuI* (New England Biolabs, Hitchin, UK) and the purified product was ligated with Adeno-X genome DNA, containing nearly the full length of the adenovirus type 5 genome lacking the E1 and E3 regions, to generate pAd.mMCP-1. Subsequently, Ad-mMCP-1 was generated by transfecting 293 cells with pAd.mMCP-1, which was linearized with *Pac*I, as described in the manual. Ad-sMCP-1 (which expresses sMCP-1), Ad-lacZ (which expresses beta-galactosidase (lacZ)) and Ad-tk (which expresses HSV-tk) were constructed as previously described and propagated in 293 cells (Figures 1b–d).[9] The rAds were purified on a cesium gradient, and the titer of rAd was determined by the 50% tissue culture infectious dose ($TCID_{50}$) method.[10]

Cell lines and culture

The mouse HCC cell lines (BNL 1NG A2, BNL 1ME A.7R. 1, MM45T.Li and Hepa 1-6) and the colon cancer cell line Colon 26 were used in these experiments. Cells were cultured in Dulbecco's modified Eagle medium (Gibco, Long Island, NY) supplemented with 10% heat-inactivated fetal bovine serum (Gibco).

ELISA for MCP-1

Aliquots of 1×10^5 HCC lines (BNL 1NG A2, BNL 1ME A.7R. 1, MM45T.Li and Hepa 1-6) and the colon cancer cell line, Colon 26 clone 20, were seeded in 1.0 ml of culture media in a six-well tissue culture plate. After 24 h, the cells were infected with Ad-mMCP-1, Ad-sMCP-1 and Ad-lacZ at various multiplicities of infection (MOI). After 48 h, the cells were harvested and sonicated to obtain the membrane fractions, and the media was collected from each well. Tumor tissues were resected on day 1 after subcutaneous injection of 5×10^6 MM45T.Li cells infected with indicated rAds (MOI 50) as described below. Tumor tissues were washed with phosphate-buffered saline (PBS) and sonicated to obtain the membrane fractions. The concentration of MCP-1 was determined by enzyme-linked immunosorbent assay (ELISA) as described previously.[11] Briefly, each well of a 96-well microtiter plate was coated with monoclonal anti-human MCP-1 antibody (ME61; 1 mg ml^{-1}) overnight at 41 °C. After washing, the plates were blocked by incubation with PBS containing 1% bovine serum albumin for 1 h at 37 °C. Diluted sample media was added, and the plate was then incubated for 2 h at 37 °C. Following incubation, the plates were washed and incubated with rabbit anti-MCP-1 antibody (1 mg ml^{-1}), followed by alkaline phosphatase-conjugated goat anti-rabbit antibody (1/12 000, Tago, Burlingame, CA), each for 2 h at 37 °C. After the plate was washed, aliquots of 1 mg ml^{-1} p-nitrophenylphosphate (Sigma, St Louis, MO) in 1 M diethanolamine (Sigma; pH 9.8) supplemented with 0.5 mM $MgCl_2$ were added to the wells, and the plate was incubated for 40 min at room temperature. After the addition of 1 M NaOH, the optical density (405 nm wavelength–OD405) was assessed by using an ELISA plate reader (MTP-120; Corona Electric, Ibaragi, Japan).

In vitro chemotaxis assay

In vitro migration assays were performed with the QCM chemotaxis cell migration assay (5 μm, Chemicon International, Temecula, CA) according to the manufacturer's instructions. Briefly, 7.5×10^4 splenocytes were resuspended in 100 μl of RPMI1640 containing 5% bovine serum albumin, and loaded into the upper well of a transwell chamber. The lower wells were filled with 150 μl of supernatant from MM45T.Li cells that were harvested 48 h after infection with rAds. The cells were incubated for 4 h at 37 °C in a humidified, 5% CO_2 atmosphere. The migrated cells were lysed and detected by the CyQuant GR dye (Molecular Probes, Eugene, OR), and fluorescence was read at an excitation wavelength of 490 nm and an emission wavelength of 520 nm in a fluorescence microplate reader (Thermo Scientific Fluoroskan Ascent FL, Thermo Fisher Scientific Oy, Vantaa, Finland).

In vitro proliferation assay

In vitro proliferation assays were performed with the CellTiter 96 Aqueous Non-Radioactive Cell Proliferation Assay (Promega, Madison, WI) according to the manufacturer's instructions. Briefly, aliquots of 1×10^4 MM45T.Li cells that were harvested 24 h after infection with rAds were seeded in a 96-well tissue culture plate and incubated for 24 h. MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium] solution was added and incubated for 2 h, and the absorbance

at 490 nm was measured by using an ELISA plate reader (MTP-120; Corona Electric).

Animal studies

The following investigations were performed in accordance with the guidelines of our Institutional Animal Care and Use Committee. Six-week-old immunocompetent female BALB/c-jcl mice (CLEA Japan, Tokyo, Japan) were injected subcutaneously on both sides of the flank on day 0 with 3×10^5 MM45T.Li cells infected with each rAd at an *in vitro* MOI of 5. For the next 5 days (days 1–5), mice received 75 mg kg^{-1} of intraperitoneally administered GCV (Tanabe Pharmaceutical Drug, Tokyo, Japan). In some series of experiment, 1 μg of the recombinant human MCP-1 in 200 μl of PBS containing 1% bovine serum albumin were injected intraperitoneally, as previously described,[12] from days 0 to 2 (3 consecutive days) in the group of the tumor cells transduced with Ad-lacZ on HSV-tk/GCV suicide therapy. Tumor sizes were measured every 3 days, and tumor volumes were calculated according to the formula (longest diameter)/(shortest diameter)^2/2.

Immunohistochemical analysis

Tumor tissues and spleens were resected on day 10. The tissue samples were embedded in OCT compound (Sakura Finetek, Toorrance, CA) and snap-frozen in liquid nitrogen. Cryostat sections of the frozen tissues were fixed with 4% paraformaldehyde in PBS, followed by washing once with distilled water for 5 min and three times with PBS for 5 min. To avoid nonspecific staining, avidin and biotin in the tissues were blocked by using a blocking kit (Vector Laboratories, Burlingame, CA).

The tissue sections were subsequently stained with rat anti-mouse CD4 Ab, rat anti-mouse CD8a Ab, rat anti-mouse CD11b Ab (BD Biosciences, San Diego, CA), or monoclonal mouse anti-human MCP-1 Ab (R&D systems, Minneapolis, MN) overnight at 4 °C. Isotype controls were also used. Then, the slides were incubated for 0.5 h at room temperature with biotinylated polyclonal rabbit anti-rat IgG (Dako Cytomation, Tokyo, Japan), or the antibodies in the M.O.M. immunodetection kit to detect mouse primary antibodies on mouse tissues (Vector Laboratories). The reactions were visualized by using a VectaStain ABC standard kit (Vector Laboratories), followed by counterstaining with hematoxylin. The positive cells were counted in 10 randomly chosen fields at 400-fold magnification by an examiner without any prior knowledge of the experimental procedures.

Quantitative real-time reverse-transcriptase PCR

Total RNA was extracted from tumor tissues resected on day 10 using an RNeasy Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. After treating the RNA preparations with ribonuclease-free DNase I (Qiagen) to remove residual DNA, cDNA was synthesized as described previously.[13] Quantitative real-time PCR was performed on a StepOne real-time PCR system (Applied Biosystems, Foster City, CA) by using the comparative C_T quantification method. TaqMan Gene Expression Assays (Applied Biosystems) containing specific primers (assay ID: tumor necrosis factor (TNF), Mm00443258_m1; glyceraldehyde-3-phosphate dehydrogenase (GAPDH), Mm99999915_g1), TaqMan MGB probe (FAM dye-labeled), and TaqMan Fast Universal PCR Master Mix were used with 10 ng cDNA to quantify the expression levels of TNF. Reactions were performed for 20 s at 95 °C followed by 40 cycles of 1 s at 95 °C and 20 s at 60 °C. The GAPDH was amplified as an internal control, and the GAPDH C_T values were subtracted from C_T values of the target genes (C_T). The ΔC_T values of tumors after immune gene therapy with both the suicide gene (HSV-tk system) and rAds were compared respectively.

Flow cytometry

MM45T.Li. cells transfected with rAds were resuspended in PBS containing 1% bovine serum albumin and 0.1% sodium azide and incubated for 30 min on ice with PE-conjugated rat anti-human MCP-1 (BD Pharmingen, San Diego, CA). The cells were washed, resuspended in PBS and analyzed in a FACScan with CellQuest software (FACSCalibur, BD Biosciences, San Jose, CA).

Depletion of macrophages/monocytes. Clodronate liposome was prepared and systemic depletion of monocytes/macrophages was performed as previously described.[14,15] Mice were intraperitoneally injected with 200 μl of clodronate liposome five times: days −2, 0, 3, 6 and 10 after tumor injection. PBS-clodronate was given in the same manner as a negative control. Depletion of CD11c-negative monocytes in blood was confirmed by flow cytometry after injection of clodronate liposome.

Statistical analysis

Mean and s.d. or s.e. were calculated for the obtained data. The statistical significance of differences between groups was evaluated by the Mann–Whitney *U*-test. $P < 0.05$ was considered statistically significant.

RESULTS

In vitro and *in vivo* MCP-1 production by cells infected with recombinant adenoviruses

When various tumor cells were infected with either Ad-mMCP-1 or Ad-sMCP-1 at an MOI of 10, the cells did not show any signs of cell death (data not shown). Both types of adenoviruses induced the secretion of human MCP-1 into the supernatants to similar levels in all the cell lines that we examined (Figure 2a). On the contrary, MCP-1 contents in the membrane fractions were higher in the cells infected with Ad-mMCP-1 than in those with Ad-sMCP-1 (Figure 2a). The proportion of MCP-1-positive cells were progressively augmented in MM45T.Li cells as the MOIs of the used Ad-mMCP-1 were increased (Figure 2b). In contract, MCP-1-positive cells were not detected in tumor cells infected with Ad-sMCP-1 and Ad-lacZ, even when the cells were infected with either vector at a MOI of 100 (Figure 2b). Thus, Ad-mMCP-1 infection can *in vitro* drive human MCP-1 expression on the cell surface, as well as its secretion into the culture medium. These results indicate that large amounts of MCP-1 protein were expressed and accumulated on the tumor cell surface when tumor cells were infected with Ad-mMCP-1 as compared with Ad-sMCP-1 *in vitro*. To define the biological functions of secreted human MCP-1 protein, we examined the migratory capacity of splenocytes to the culture supernatants obtained 24 h after the infection. The supernatants from either Ad-mMCP-1- or Ad-sMCP-1-infected cells enhanced the transmigration of splenic lymphocytes to similar extents, compared with those from Ad-lacZ-infected cells (Figure 2c). These results indicated that biologically active human MCP-1 was secreted into the culture supernatants.

Proliferation of tumor cells infected with rAds *in vitro* and *in vivo*

To quantify the proliferation of tumor cells infected with rAds, the MTS assay was performed 24 h after infection. The optical absorbance at 490 nm of tumor cells did not change in the presence or absence of rAd infection (Figure 3a). Next, tumor cells infected with rAds (MOI 10) *ex vivo* were transferred subcutaneously in syngeneic wild-type mice, and tumor development was monitored (Figure 3b). Tumor cells infected with Ad-mMCP-1 and Ad-sMCP-1 showed similar growth rates to tumor cells infected with Ad-lacZ. These results indicate that infection with the rAds used in this study did not affect the proliferation of tumor cells *in vitro* and *in vivo*, and that the delivery of mMCP-1 did not display antitumor activity when used alone. In addition, the levels of MCP-1 expression were confirmed immunohistochemically in the subcutaneous tumor tissues (Figure 3c). MCP-1-positive tumor cells were detected in tumor tissues infected with Ad-mMCP-1; however, the cells were negative for MCP-1 staining in tissues infected with Ad-sMCP-1 and Ad-lacZ. Moreover, larger amounts of MCP-1 were detected in the tumor tissues of the mice treated with Ad-mMCP-1 than in those with Ad-sMCP-1 (Figure 3d). The data indicated that large amounts of MCP-1 protein were expressed and accumulated on the tumor cell surface when the tumor cells were infected with Ad-mMCP-1 *in vivo*.

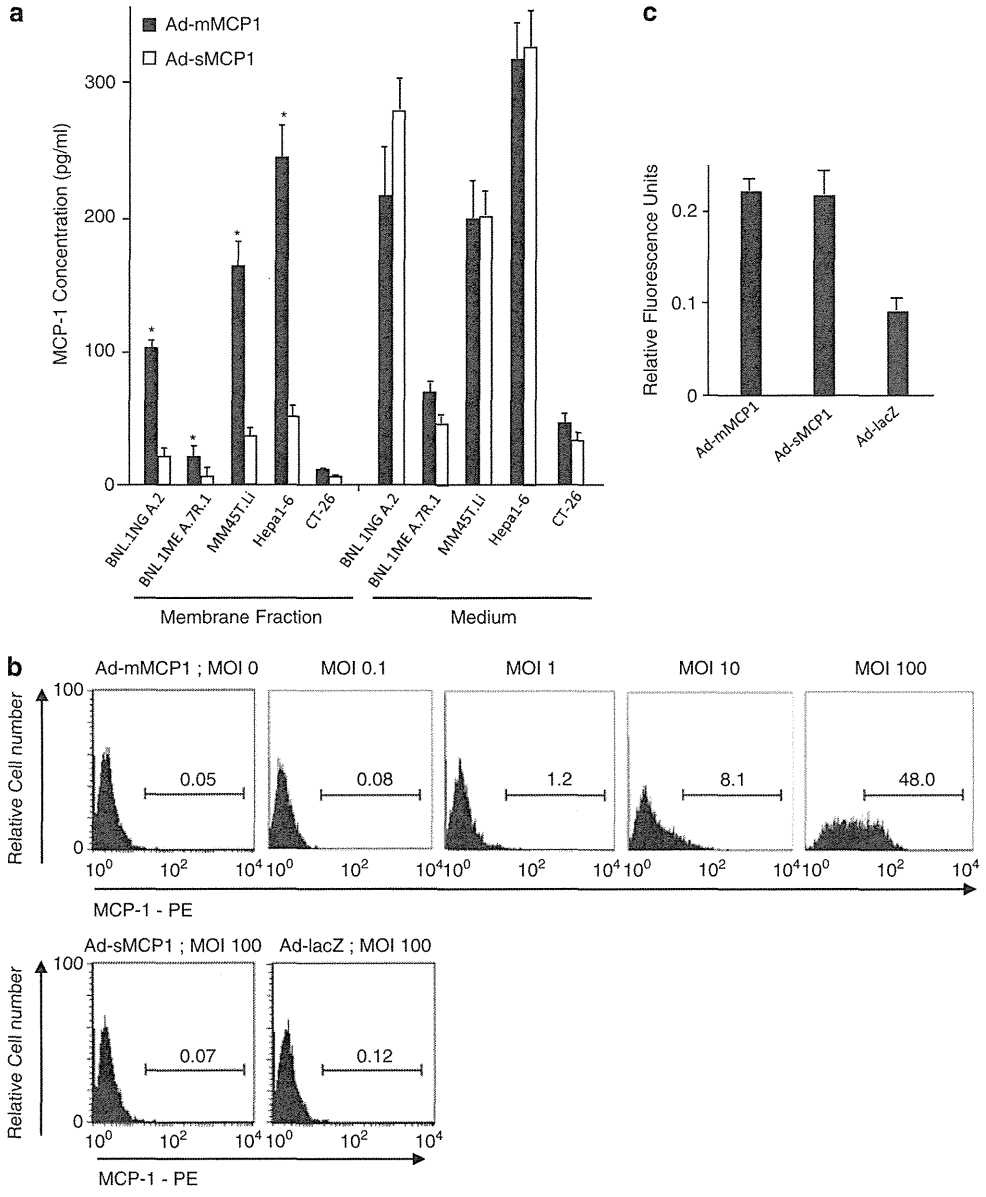

Figure 2. Monocyte chemoattractant protein-1 (MCP-1) production in tumor cells infected with recombinant adenovirus vectors (rAds). (**a**) Concentrations of MCP-1 in the membrane fractions and in the media of hepatoma cells (BNL 1NG A.2, BNL 1ME A.7R.1, MM45T.Li and Hepa1-6) and colon cancer cells (colon 26 clone 20) infected with rAds at multiplicities of infection (MOI) of 10 were measured by enzyme-linked immunosorbent assay (ELISA). Each value is the mean s.d. of triplicate experiments. *, $P<0.05$ when compared with Ad-soluble MCP-1 (sMCP-1) by the Mann–Whitney's *U*-test. (**b**) Surface expression of MCP-1 on MM45T.Li cells infected with rAds (Ad-mMCP-1, Ad-sMCP-1 and Ad-lacZ) at the indicated MOIs was assessed by flow cytometry by using PE-conjugated anti-human MCP-1 antibody. Histograms represent MCP-1 staining of tumor cells. Numbers indicate percentages of MCP-1-positive cells. Surface MCP-1-positive cells were detected in 0.08, 1.2, 8.1 and 48.0% of MM45T.Li cells infected by Ad-mMCP-1 at MOIs of 0.1, 1, 10 and 100, respectively. The results are representative of three independent experiments. (**c**) The migratory activity of MCP-1 secreted from rAd-infected tumor cells. Mouse splenic lymphocytes were loaded into the upper wells of transwell chambers, and supernatants of tumor cells infected with rAds at an MOI of 10 were added to the lower wells. Cells that migrated through the 8-μm pores to the feeder tray after 4 h incubation were lysed and detected by CyQuant GR dye that exhibits enhanced fluorescence upon binding cellular nucleic acids. Each value is the mean s.e. of data from three separate migration chambers.

Potentiation of HSV-tk/GCV suicide therapy by co-infection with Ad-mMCP-1

We previously demonstrated that the gene delivery of Ad-sMCP-1 enhanced the antitumor effects of the HSV-tk/GCV system.[4–6,16,17] Hence, we compared the effects of Ad-mMCP-1 and Ad-sMCP-1 infection on HSV-tk/GCV suicide therapy. When MM45T.Li cells were co-infected with Ad-tk and Ad-lacZ, and received GCV, tumor growth was delayed marginally but not significantly (Figure 4a). Co-infection with Ad-tk and Ad-sMCP-1 retarded tumor growth significantly after GCV treatment. Moreover, tumor growth was almost abrogated by the combination of co-infection with Ad-tk and Ad-mMCP-1, and GCV treatment. To address whether the antitumor effects could be induced not only by the gene delivery using Ad vector, but also by the administration of recombinant protein, we gave intraperitoneally recombinant MCP-1 to the animals, which were injected with tumor cells, treated with

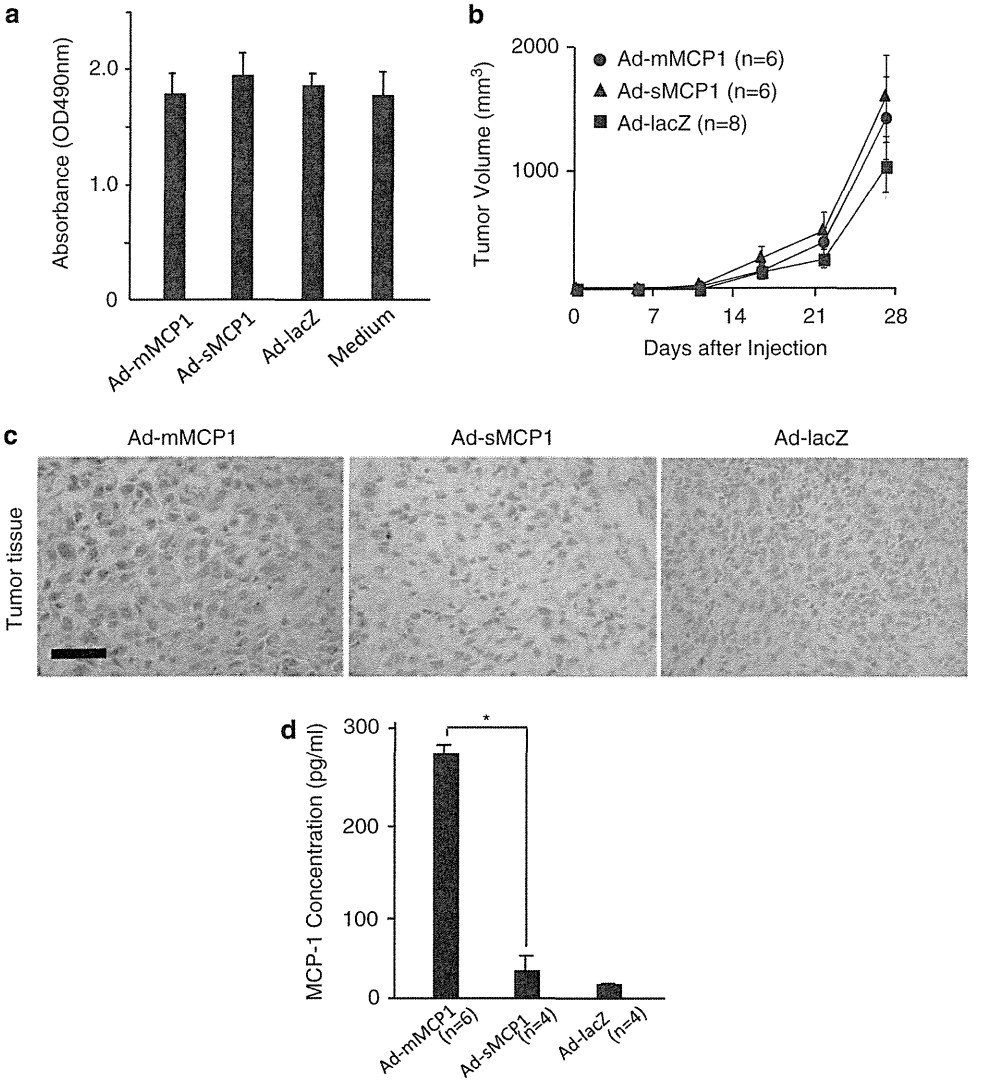

Figure 3. Proliferation of tumor cells infected with recombinant adenovirus vectors (rAds) *in vitro* and *in vivo*. (**a**) Tumor cell growth after the infection of indicated rAds *in vitro*. A total of 5×10^5 of MM45T.Li cells were infected with the rAds at multiplicities of infection (MOI) of 10 and incubated for 24 h. The cell numbers were quantitated by MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) assay. The absorbance was determined at 490 nm with a microplate reader. Each value is the mean s.d. of data from triplicate experiments. (**b**) Tumor cell growth after infection of indicated rAds in mice. BALB/c mice were subcutaneously injected with 3×10^5 MM45T.Li cells infected with the rAds at an MOI of 10 on day 0. Tumor diameters were monitored. Each value is the mean s.e. (**c**) Immunohistochemical analysis of subcutaneous tumor tissues 7 days after the injection in panel **b**. Tissues were stained and visualized by using anti-human monocyte chemoattractant protein-1 (MCP-1) Ab and ABC methods. MCP-1 expression was seen as brown in the cytoplasm of tumor cells. The bar represents 30 μm. Original magnification, ×400. (**d**) Concentrations of MCP-1 were measured in the s.c. tumor tissues resected on day 1 after injection of 5×10^6 MM45T.Li cells infected with the indicated rAds at an MOI of 50 by enzyme-linked immunosorbent assay (ELISA). Each value is the mean s.d. of duplicate experiments. $^*P < 0.05$ when compared with Ad-sMCP-1 by the Mann–Whitney's *U*-test.

Ad-lacZ and Ad-tk. The systemic administration of recombinant MCP-1 rather enhanced tumor growth (Figure 4b). As we previously demonstrated that MCP-1 can promote tumor growth in a context-dependent manner by recruiting macrophages, which can secrete an angiogenic factor, the vascular endothelial growth factor,[16,18] systemic MCP-1 injection may promote tumor growth by its pro-angiogenic activities.

Recruitment and activation of macrophages and T lymphocytes in tumor tissues

The GCV treatment following co-infection with Ad-lacZ and Ad-tk failed to increase the intratumoral numbers of Mac-1-positive macrophages, CD4-positive lymphocytes and CD8-positve lymphocytes, compared with GCV treatment following Ad-lacZ infection (Figures 5A and B). GCV administration following co-infection with Ad-sMCP-1 and Ad-tk increased the intratumoral Mac-1-positive macrophage, CD4-positive lymphocyte and CD8-positve lymphocyte numbers (Figure 5A and B). The increases were further enhanced by GCV treatment following co-infection with Ad-mMCP-1 and Ad-tk (Figures 5A and B). Moreover, intratumoral mRNA expression of TNF, a known macrophage and T lymphocyte secretagogue, was markedly enhanced in tumors co-infected with Ad-mMCP-1, compared with those with Ad-sMCP-1 plus Ad-tk. To evaluate the functional contribution of intratumoral immune cells, we depleted CD11c-negative

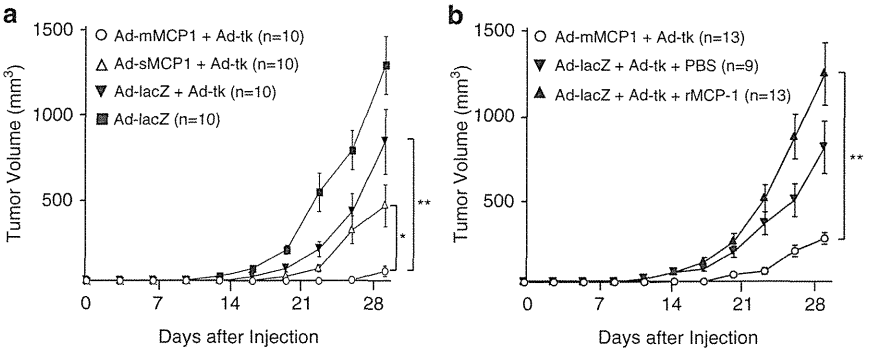

Figure 4. Antitumor effects of recombinant adenovirus vector (rAds) *in vivo*. (**a**) BALB/c mice were subcutaneously injected with 3×10^5 MM45T.Li cells infected with rAds Ad-membrane-bound monocyte chemoattractant protein-1 (mMCP-1) + Ad-tk, Ad-soluble MCP-1 (sMCP-1) + Ad-tk, Ad-lacZ + Ad-tk, and Ad-lacZ at multiplicities of infection (MOI) of 10 on day 0. Subsequently, 75 mg kg^{-1} of ganciclovir (GCV) was administered for 5 consecutive days (days 2–6). Each value is the mean s.e. of triplicate experiments. $^*P < 0.05$ when compared with Ad-sMCP-1 + Ad-tk, and $^{**}P < 0.01$ when compared with Ad-lacZ + Ad-tk by the Mann–Whitney's *U*-test. (**b**) BALB/c mice were injected with MM45T.Li cells and treated as described in the legend to (**a**). In Ad-lacZ + Ad-tk + recombinant human MCP-1 (rMCP-1) group, the mice were injected with 1 μg of rMCP-1 intraperitoneally from days 0 to 2 (3 consecutive days). The mice were injected with phosphate-buffered saline (PBS) as controls. Tumor sizes were measured every 3 days. Each value is the mean s.e. $^{**}P < 0.01$ when compared with Ad-lacZ + Ad-tk + rMCP-1 by the Mann–Whitney's *U*-test.

monocytes/macrophages by intraperitoneal administration of clodronate liposome in the current mouse model. The monocyte/macrophage-depleted mice developed larger tumor than those treated with PBS liposome (Figure 5D), indicating that monocytes/macrophages were critically involved in the suppression of tumor growth by Ad-mMCP-1. Collectively, these data demonstrate that the delivery of mMCP-1 promoted the recruitment and activation of macrophages and T lymphocytes in tumor tissues, presumably leading to the beneficial antitumor responses in this model.

DISCUSSION

We have proposed a strategy for improving the efficacy of suicide gene-based gene therapy by the combined heterochronic administration of *HSV-tk* and *MCP-1* genes.[4–6,16–18] In the current study, we generated recombinant adenovirus Ad-mMCP-1 expressing a fusion protein containing the human MCP-1 cDNA fused with the membrane-spanning domain of fractalkine/CX3CL1, to drive more efficient and sessile expression of MCP-1. Ad-mMCP-1 infection did not affect the proliferation of MM45T.Li tumor cells *in vitro* or *in vivo*, by itself. Of interest is that Ad-mMCP-1 infection pontentiated HSV-tk/GCV suicide therapy more efficiently than Ad-sMCP-1. Moreover, Ad-mMCP-1-mediated antitumor effects were associated with the recruitment and activation of macrophages and T lymphocytes in tumor tissues. Collectively, the delivery of membrane-bound *MCP-1* gene can augment antitumor effects caused by the HSV-tk/GCV system in an immunecompetent mouse model of liver tumor and therefore, can be a novel strategy of immune-based gene therapy to prevent tumor proliferation and recurrence in patients with HCC.

Chimeric membrane-bound cytokine gene expression vectors were generated to drive the efficient expression on tumor cell surface and to reduce the severe side effects caused by systemic administration of high doses of cytokines. With this maneuver, cytokines can be anchored on the cell plasma membrane. As a consequence, a locally high concentration of cytokines can be achieved with ease and their *in vivo* half-life can be prolonged in the tumor site. The availability of cytokines on tumor cell surface can eventually bring immune cells to the tumor site for better antigen uptake and stimulation, thereby inducing antitumor effects at a higher efficiency. On the basis of these assumptions, this type of modified cytokine gene therapy has been reported on interleukin-2,[13,19,20] interleukin-12,[21] fractalkine (CX3CL1)[22] and TNF.[23] Indeed, accumulating evidence revealed that the membrane-bound form of cytokine genes can exhibit more antitumor effects than the corresponding soluble ones. Likewise, the current study confirms that the membrane-bound form of MCP-1 can attract more immune cells, including monocytes/macrophages and T lymphocytes, to the tumor sites and can induce the expression of TNF. In addition, TNF can activate endothelial cells to express several adhesion molecules, such as the intercellular adhesion molecules and vascular cell adhesion molecules.[24–27] Circulating immune cells can utilize these adhesion molecules to effectively transmigrate into the tissues in addition to the direct chemotactic activities exerted by MCP-1.

The effects of MCP-1 on tumor growth was controversial, either destructive[28–30] or protective[31,32] in a context-dependent manner. Likewise, murine colon carcinoma cell expressing MCP-1 failed to metastasize when injected into mice,[28] whereas other carcinoma cells showed enhanced metastasis.[31] These discrepancies may be explained by the observations reported by Nesbit *et al.*[33] They demonstrated that low-level MCP-1 secretion with modest monocyte infiltration resulted in tumor formation, whereas high secretion was associated with massive monocyte/macrophage infiltration into the tumor mass, leading to its destruction within a few days. Thus, systemic MCP-1 administration may not be able to induce massive monocyte/macrophage infiltration into tumor mass and may promote tumor mass as we observed in the present study. Moreover, we previously revealed that suicide therapy can induce tumor cell apoptosis and that apoptotic tumor cells can secrete MCP-1 more efficiently, thereby recruiting a massive number of macrophages and retarding tumor growth.[4] Consistently, we further demonstrated that the delivery of an optimal amount of rAd expressing MCP-1 enhanced the antitumor effects of the HSV-tk/GCV system in a model of HCC.[16,18] Infection with Ad-mMCP-1 can sustain MCP-1 expression in tumor tissues more efficiently than that with Ad-sMCP-1 as evidenced by an immunohistochemical analysis on the infected tumor tissues. The sustained MCP-1 expression can potentiate suicide gene therapy more efficiently.

The present data suggest that the use of Ad-mMCP-1 can be promising, but several problems remain to be solved before the clinical application. First, subcutaneous tumor models of an HCC cell line may not be relevant to HCCs in patients. However, in cases of nonsurgical procedures for HCC treatment in patients, such as percutaneous radiofrequency ablation therapy[34] and

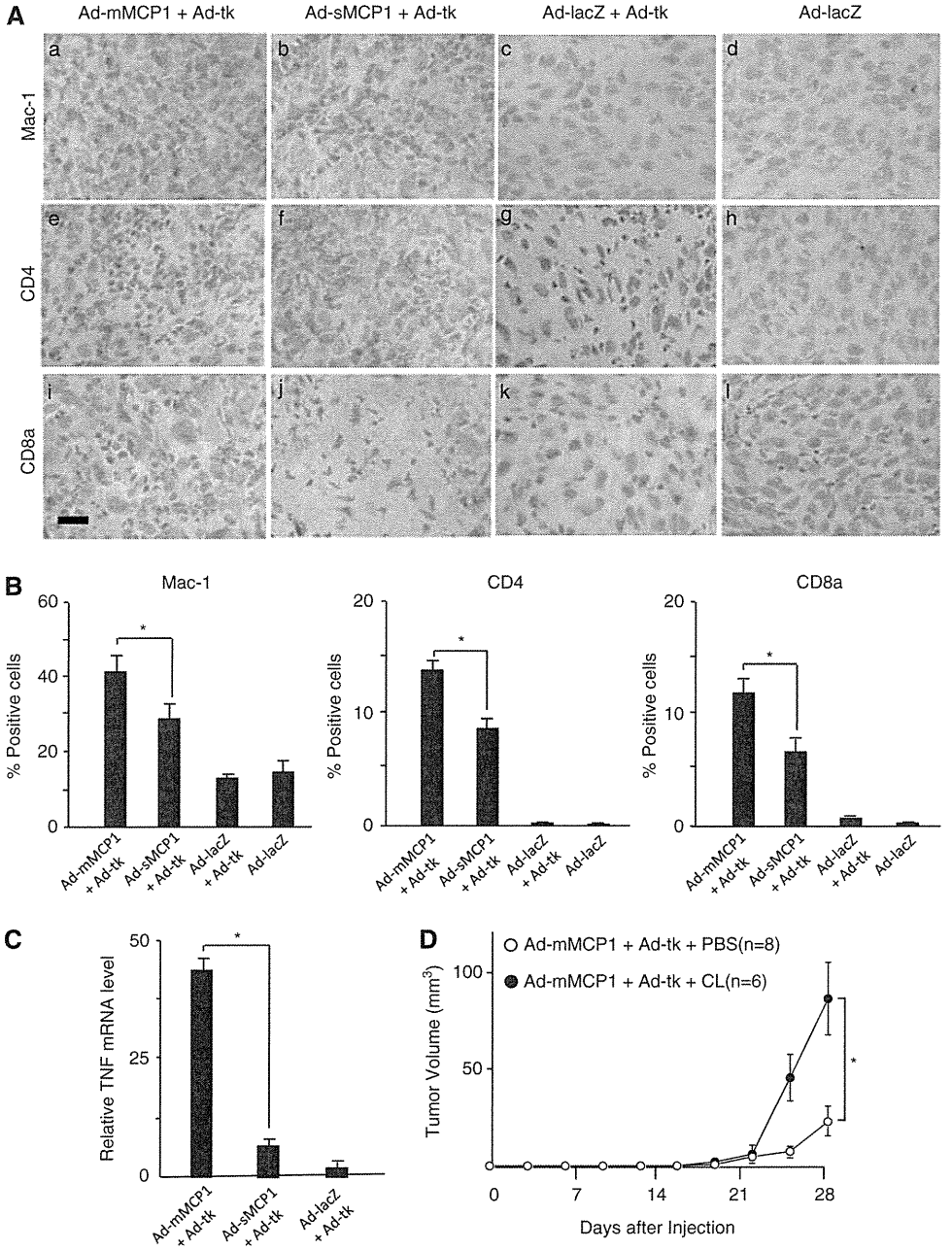

Figure 5. Immunohistochemical analysis for Mac-1-, CD4- and CD8a-positive cells and expression of tumor necrosis factor (TNF) mRNA in tumor tissues. In the experiment described in the legend to Figure 4, tumors were resected on day 10 and used for the analyses. (**A**) Tumor tissues treated with Ad-membrane-bound monocyte chemoattractant protein-1 (mMCP-1) + Ad-tk (a, e and i), Ad-soluble MCP-1 (sMCP-1) + Ad-tk (b, f and j), Ad-lacZ + Ad-tk (c, g and k) and Ad-lacZ (d, h and l) were stained with anti-Mac-1 antibody (a–d), anti-CD4 antibody (e–h) or anti-CD8a antibody (i–l). Positive cells are seen as brown. Original magnification, × 400. (**B**) Quantitative morphometric analysis showing the proportions of positive cells in areas of 100 tumor cells. Ten high-power (× 400) fields of tumor tissue were examined. Results are expressed as means per 1000 hepatoma cells. Values are the means s.d. of triplicate experiments. *$P<0.05$ when compared with Ad-sMCP-1 by the Mann–Whitney's *U*-test. (**C**) Real-time PCR analysis for TNF mRNA expression in tumor tissues, presented relative to glyceraldehyde 3-phosphate dehydrogenase (GAPDH) mRNA. Each value is the mean s.d. of duplicate experiments. *$P<0.05$ when compared with Ad-sMCP-1 + Ad-tk by the Mann–Whitney *U*-test. (**D**) BALB/c mice were subcutaneously injected with MM45T.Li cells infected with indicated rAds and treated as described in the legend to Figure 4a. In the Ad-mMCP-1 + Ad-tk + clodronate liposome (CL) group, mice were injected with 200 μl of CL to deplete monocytes/macrophages as described in Materials and methods. The mice were injected with phosphate-buffered saline (PBS) liposome as controls. Tumor sizes were measured every 3 days. Each value is the mean s.e. *$P<0.05$ when compared with Ad-mMCP-1 + Ad-tk + PBS by the Mann–Whitney's *U*-test.

transcatheter arterial chemotherapy,[35] administration of the current rAd vectors could be easily applied, immediately once the standard nonsurgical procedures to ensure tumor cell killing. Moreover, rAd can elicit its immunogenicity or cytotoxicity when administered in HCC patients, particularly by the use of intraarterial procedures. Actually, the infection of high doses of rAds causes severe unexpected side effects.[36] However, the delivery of membrane-bound form of the *MCP-1* gene can cause

a high and effective concentration at the tumor sites even when it is administered at a relatively lower titer, and therefore, can evade severe adverse effects caused frequently by the administration of a high titer of adenovirus vectors.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Mariko Katsuda for assistance with histopathological analysis and immunohistochemistry, and Maki Kawamura and Chiharu Minami for providing animal care.

REFERENCES

1 Venook AP. Treatment of hepatocellular carcinoma: too many options? *J Clin Oncol* 1994; **12**: 1323-1334.
2 Trinchet JC, Beaugrand M. Treatment of hepatocellular carcinoma in patients with cirrhosis. *J Hepatol* 1997; **27**: 756-765.
3 Bruix J. Treatment of hepatocellular carcinoma. *Hepatology* 1997; **25**: 259-262.
4 Sakai Y, Kaneko S, Nakamoto Y, Kagaya T, Mukaida N, Kobayashi K. Enhanced antitumor effects of herpes simplex virus thymidine kinase/ganciclovir system by codelivering monocyte chemoattractant protein-1 in hepatocellular carcinoma. *Cancer Gene Ther* 2001; **8**: 695-704.
5 Tsuchiyama T, Kaneko S, Nakamoto Y, Sakai Y, Honda M, Mukaida N *et al.* Enhanced antitumor effects of a bicistronic adenovirus vector expressing both herpes simplex virus thymidine kinase and monocyte chemoattractant protein-1 against hepatocellular carcinoma. *Cancer Gene Ther* 2003; **10**: 260-269.
6 Kagaya T, Nakamoto Y, Sakai Y, Tsuchiyama T, Yagita H, Mukaida N *et al.* Monocyte chemoattractant protein-1 gene delivery enhances antitumor effects of herpes simplex virus thymidine kinase/ganciclovir system in a model of colon cancer. *Cancer Gene Ther* 2006; **13**: 357-366.
7 Freund CT, Sutton MA, Dang T, Contant CF, Rowley D, Lerner SP. Adenovirus-mediated combination suicide and cytokine gene therapy for bladder cancer. *Anticancer Res* 2000; **20**: 1359-1365.
8 Kim YS. Tumor therapy applying membrane-bound form of cytokines. *Immune Netw* 2009; **9**: 158-168.
9 Sato Y, Tanaka K, Lee G, Kanegae Y, Sakai Y, Kaneko S *et al.* Enhanced and specific gene expression via tissue-specific production of Cre recombinase using adenovirus vector. *Biochem Biophys Res Commun* 1998; **244**: 455-462.
10 Kanegae Y, Makimura M, Saito I. A simple and efficient method for purification of infectious recombinant adenovirus. *Jpn J Med Sci Biol* 1994; **47**: 157-166.
11 Ko Y, Mukaida N, Panyutich A, Voitenok NN, Matsushima K, Kawai T *et al.* A sensitive enzyme-linked immunosorbent assay for human interleukin-8. *J Immunol Methods* 1992; **149**: 227-235.
12 Nakano Y, Kasahara T, Mukaida N, Ko YC, Nakano M, Matsushima K. Protection against lethal bacterial infection in mice by monocyte-chemotactic and -activating factor. *Infect Immun* 1994; **62**: 377-383.
13 Ji J, Li J, Holmes LM, Burgin KE, Yu X, Wagner TE *et al.* Glycoinositol phospholipid-anchored interleukin 2 but not secreted interleukin 2 inhibits melanoma tumor growth in mice. *Mol Cancer Ther* 2002; **1**: 1019-1024.
14 Lu P, Li L, Liu G, van Rooijen N, Mukaida N, Zhang X. Opposite roles of CCR2 and CX3CR1 macrophages in alkali-induced corneal neovascularization. *Cornea* 2009; **28**: 562-569.
15 Sadahira Y, Yasuda T, Yoshino T, Manabe T, Takeishi T, Kobayashi Y *et al.* Impaired splenic erythropoiesis in phlebotomized mice injected with CL2MDP-liposome: an experimental model for studying the role of stromal macrophages in erythropoiesis. *J Leukoc Biol* 2000; **68**: 464-470.
16 Tsuchiyama T, Nakamoto Y, Sakai Y, Mukaida N, Kaneko S. Optimal amount of monocyte chemoattractant protein-1 enhances antitumor effects of suicide gene therapy against hepatocellular carcinoma by M1 macrophage activation. *Cancer Sci* 2008; **99**: 2075-2082.
17 Tsuchiyama T, Nakamoto Y, Sakai Y, Marukawa Y, Kitahara M, Mukaida N *et al.* Prolonged, NK cell-mediated antitumor effects of suicide gene therapy combined with monocyte chemoattractant protein-1 against hepatocellular carcinoma. *J Immunol* 2007; **178**: 574-583.
18 Kakinoki K, Nakamoto Y, Kagaya T, Tsuchiyama T, Sakai Y, Nakahama T *et al.* Prevention of intrahepatic metastasis of liver cancer by suicide gene therapy and chemokine ligand 2/monocyte chemoattractant protein-1 delivery in mice. *J Gene Med* 2010; **12**: 1002-1013.
19 Chang MR, Lee WH, Choi JW, Park SO, Paik SG, Kim YS. Antitumor immunity induced by tumor cells engineered to express a membrane-bound form of IL-2. *Exp Mol Med* 2005; **37**: 240-249.
20 Ji J, Li J, Holmes LM, Burgin KE, Yu X, Wagner TE *et al.* Synergistic anti-tumor effect of glycosylphosphatidylinositol-anchored IL-2 and IL-12. *J Gene Med* 2004; **6**: 777-785.
21 Nagarajan S, Selvaraj P. Glycolipid-anchored IL-12 expressed on tumor cell surface induces antitumor immune response. *Cancer Res* 2002; **62**: 2869-2874.
22 Tang L, Hu HD, Hu P, Lan YH, Peng ML, Chen M *et al.* Gene therapy with CX3CL1/Fractalkine induces antitumor immunity to regress effectively mouse hepatocellular carcinoma. *Gene Ther* 2007; **14**: 1226-1234.
23 Rieger R, Whitacre D, Cantwell MJ, Prussak C, Kipps TJ. Chimeric form of tumor necrosis factor-alpha has enhanced surface expression and antitumor activity. *Cancer Gene Ther* 2009; **16**: 53-64.
24 Nooijen PT, Eggermont AM, Verbeek MM, Schalkwijk L, Buurman WA, de Waal RM *et al.* Transient induction of E-selectin expression following TNF alpha-based isolated limb perfusion in melanoma and sarcoma patients is not tumor specific. *J Immunother Emphasis Tumor Immunol* 1996; **19**: 33-44.
25 Yang L, Froio RM, Sciuto TE, Dvorak AM, Alon R, Luscinskas FW. ICAM-1 regulates neutrophil adhesion and transcellular migration of TNF-alpha-activated vascular endothelium under flow. *Blood* 2005; **106**: 584-592.
26 Vanhee D, Delneste Y, Lassalle P, Gosset P, Joseph M, Tonnel AB. Modulation of endothelial cell adhesion molecule expression in a situation of chronic inflammatory stimulation. *Cell Immunol* 1994; **155**: 446-456.
27 VandenBerg E, Reid MD, Edwards JD, Davis HW. The role of the cytoskeleton in cellular adhesion molecule expression in tumor necrosis factor-stimulated endothelial cells. *J Cell Biochem* 2004; **91**: 926-937.
28 Huang S, Singh RK, Xie K, Gutman M, Berry KK, Bucana CD *et al.* Expression of the JE/MCP-1 gene suppresses metastatic potential in murine colon carcinoma cells. *Cancer Immunol Immunother* 1994; **39**: 231-238.
29 Rollins BJ, Sunday ME. Suppression of tumor formation *in vivo* by expression of the JE gene in malignant cells. *Mol Cell Biol* 1991; **11**: 3125-3131.
30 Nokihara H, Yanagawa H, Nishioka Y, Yano S, Mukaida N, Matsushima K *et al.* Natural killer cell-dependent suppression of systemic spread of human lung adenocarcinoma cells by monocyte chemoattractant protein-1 gene transfection in severe combined immunodeficient mice. *Cancer Res* 2000; **60**: 7002-7007.
31 Nakashima E, Mukaida N, Kubota Y, Kuno K, Yasumoto K, Ichimura F *et al.* Human MCAF gene transfer enhances the metastatic capacity of a mouse cachectic adenocarcinoma cell line *in vivo*. *Pharm Res* 1995; **12**: 1598-1604.
32 Ueno T, Toi M, Saji H, Muta M, Bando H, Kuroi K *et al.* Significance of macrophage chemoattractant protein-1 in macrophage recruitment, angiogenesis, and survival in human breast cancer. *Clin Cancer Res* 2000; **6**: 3282-3289.
33 Nesbit M, Schaider H, Miller TH, Herlyn M. Low-level monocyte chemoattractant protein-1 stimulation of monocytes leads to tumor formation in nontumorigenic melanoma cells. *J Immunol* 2001; **166**: 6483-6490.
34 Curley SA. Radiofrequency ablation of malignant liver tumors. *Ann Surg Oncol* 2003; **10**: 338-347.
35 Tung-Ping Poon R, Fan ST, Wong J. Risk factors, prevention, and management of postoperative recurrence after resection of hepatocellular carcinoma. *Ann Surg* 2000; **232**: 10-24.
36 Marshall E. Gene therapy death prompts review of adenovirus vector. *Science* 1999; **286**: 2244-2245.

Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns

Wei Qu,[1,5] Shin-ichi Hashimoto,[1] Atsuko Shimada,[2] Yoichiro Nakatani,[1] Kazuki Ichikawa,[1] Taro L. Saito,[1] Katsumi Ogoshi,[3] Kouji Matsushima,[3] Yutaka Suzuki,[4] Sumio Sugano,[4] Hiroyuki Takeda,[2] and Shinichi Morishita[1,5]

[1]*Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa 277-0882, Japan;* [2]*Department of Biological Sciences, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan;* [3]*Department of Molecular Preventive Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan;* [4]*Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo 108-8639, Japan*

5-methyl-cytosines at CpG sites frequently mutate into thymines, accounting for a large proportion of spontaneous point mutations. The repair system would leave substantial numbers of errors in neighboring regions if the synthesis of erased gaps around deaminated 5-methyl-cytosines is error-prone. Indeed, we identified an unexpected genome-wide role of the CpG methylation state as a major determinant of proximal natural genetic variation. Specifically, 507 Mbp (~18%) of the human genome was within 10 bp of a CpG site; in these regions, the single nucleotide polymorphism (SNP) rate significantly increased by ~50% ($P < 10^{-566}$ by a two-proportion *z*-test) if the neighboring CpG sites are methylated. To reconfirm this finding in another vertebrate, we compared six single-base resolution methylomes in two inbred medaka (*Oryzias latipes*) strains with sufficient genetic divergence (3.4%). We found that the SNP rate also increased by ~50% ($P < 10^{-2170}$), and the substitution rates in all dinucleotides increased simultaneously ($P < 10^{-441}$) around methylated CpG sites. In the hypomethylated regions, the "CGCG" motif was significantly enriched ($P < 10^{-680}$) and evolutionarily conserved ($P = \sim 0.203\%$), and slow CpG deamination rather than fast CpG gain was seen, indicating a possible role of CGCG as a candidate *cis*-element for the hypomethylation state. In regions that were hypermethylated in germline-like tissues but were hypomethylated in somatic liver cells, the SNP rate was significantly smaller than that in hypomethylated regions in both tissues, suggesting a positive selective pressure during DNA methylation reprogramming. This is the first report of findings showing that the CpG methylation state is significantly correlated with the characteristics of evolutionary change in neighboring DNA.

[Supplemental material is available for this article.]

Our understanding of the role of DNA methylation in genetic variation is limited to the observation that methylated cytosines at CpG sites mutate to thymines at very high frequencies (Lindahl and Nyberg 1972; Coulondre et al. 1978; Cooper and Krawczak 1989). This has been confirmed in genome-wide analyses (Sved and Bird 1990; Lander et al. 2001; Venter et al. 2001). Recently, precise methylation maps at single-base resolution have been created using next-generation DNA sequencing technology, confirming the overabundance of G:C→A:T transitions (Ossowski et al. 2010). In contrast, hypomethylated cytosine nucleotides seem to escape mutation; e.g., comparative genome analysis of human and primates genomes showed a lower CpG mutation rate in CpG-rich promoters that are mostly hypomethylated in the germline (Saxonov et al. 2006; Weber et al. 2007). Although the system for repairing mutations of 5-methyl cytosines to thymines is not well understood, it must be relatively ineffective because a large proportion of spontaneous point mutations are C-to-T mutations. The erroneous repair system may also leave substantial errors in neighboring regions if the synthesis of erased gaps around deaminated 5-methyl-cytosines is error-prone. The genome-wide methylation information provides an unprecedented opportunity to examine whether the CpG methylation state affects proximal genetic variation.

A related problem is the allele-specific DNA methylation (ASM), different methylation patterns in parental alleles (Chandler et al. 1987). ASM is involved in genomic imprinting and X-inactivation in females to achieve dosage compensation. Autosomal ASM has also been reported in the human and mouse genomes (Yamada et al. 2004; Heijmans et al. 2007; Zhang et al. 2009), and recent genome-wide collections of DNA methylation states have started uncovering that ASM is prevalent in the mammalian genome (Kerkel et al. 2008; Hellman and Chess 2010; Schalkwyk et al. 2010). Genome-wide studies indicate that ASM is associated with *cis*-acting polymorphisms (local genotypes), such as mutations in CpG sites (Schilling et al. 2009; Shoemaker et al. 2010), suggesting Mendelian inheritance patterns of ASM. A number of genomic regions of intermediate DNA methylation level are found in the human genome (Deng et al. 2009; Lister et al. 2009) and are thought to be largely a consequence of ASM and to have implications for complex disease genetics (Meaburn et al. 2010). However, little is known about fundamental *cis*-elements for inducing DNA methylation. Sequence motifs that are significantly enriched and evolutionarily conserved in hypomethylated regions could be candidate *cis*-elements; however, it is highly nontrivial to identify conserved sequence motifs due to the quite low incidence of genetic variation (~0.1%) in the human and mouse genomes. We approach this problem through the study of correlation between CpG methylation state and proximal genetic variation in the medaka genome.

[5]Corresponding authors
E-mail quwei@cb.k.u-tokyo.ac.jp
E-mail moris@cb.k.u-tokyo.ac.jp
Article published online before print. Article, supplemental material, and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.140236.112. Freely available online through the *Genome Research* Open Access option.

Results

We examined the relationship between DNA methylation and genetic variation in the human genome using publicly available data. Specifically, we examined the methylome data from human germline sperm cells (Molaro et al. 2011) and the reference single nucleotide polymorphisms (refSNPs) of the CEU population collected by the HapMap Project (The International HapMap Consortium 2003). Although substantial data are available on the human genome, the incidence of genetic variation in the human genome is quite low (~0.1%), which is not sufficient to perform a detailed analysis on the dinucleotide patterns of genetic variations associated with hyper- and hypomethylated regions. For this purpose, the medaka (*Oryzias latipes*) model system provides an ideal resource in vertebrates because two medaka inbred strains, Hd-rR and HNI, have a sufficiently high incidence of genetic variation (~3.4%) (Kasahara et al. 2007). The Hd-rR and HNI medaka inbred strains originated in the southern and northern parts of Japan, respectively, which were separated by a watershed, and the strains diverged ~18 million years ago (Setiamarga et al. 2009). Another merit of the medaka is its abundance of germline cells in the testes and blastulae (half-day embryos, in which some of the cells remain totipotent [Hong et al. 1998]), which provides information on the primary relationships between the methylation pattern and genomic variation during the course of evolution. Thus, we newly generated six methylomes of medaka genomes at single-base resolution using genomic DNA from three libraries (liver, blastulae, and testes) of the two strains, treating it with bisulphite, and subjecting it to Illumina sequencing. In total, 1.8 billion reads were produced, and 44 billion nucleotides were uniquely mapped to the genome, with an average depth of ninefold for each cell line of each strain (Supplemental Table S1). An average of ~78% of each medaka genome was covered by at least one read (Supplemental Table S1). The bisulphite conversion procedure was highly effective, as validated by a control experiment using yeast (*Saccharomyces cerevisiae* S288C) (Supplemental Table S2). In yeast, cytosine nucleotides are not methylated, and indeed, none of yeast cytosine nucleotides was uncovered to be hypermethylated (methylation level $\geq$ 0.8) (Supplemental Fig. S1). Data for the entire methylome of medaka at single base resolution can be accessed at http://utgenome.org/medaka/methylome/.

Substantial cytosines at non-CpG sites are methylated in human ES and *Arabidopsis* (Cokus et al. 2008; Lister et al. 2008, 2009); however, methylation at non-CpG sites was very rare in human sperm (Molaro et al. 2011) and in all of the three medaka tissues (Supplemental Fig. S2); i.e., few non-CpG cytosines (~0.02%) were observed to be hypermethylated. To examine whether cytosine methylation state, unmethylated or methylated, affects proximal genetic variation, both unmethylated and methylated cytosines should be sufficiently available. In CpG sites, the ratio of unmethylated cytosines:methylated cytosines was ~10%:~90%, showing the abundance of unmethylated cytosines. In non-CpG sites, however, the ratio was 99.98%:0.02%, indicating the lack of methylated cytosines. Thus, we focused on methylation patterns of CpG sites in subsequent analyses. The distribution of methylation levels was found to be bimodal, and the vast majority of CpG sites were highly methylated in human sperm (Molaro et al. 2011) and in each medaka library (Supplemental Fig. S2). CpG islands may be appropriate regions in which to measure the effect of methylation pattern on the incidence of genetic variation; however, they occupy a relatively small proportion of the human genome (~1%) as well as the medaka genome (~3%) and exist mostly in CpG-rich regions. Therefore, regions with fewer CpG sites are overlooked. For defining more comprehensive characteristics of CpG sites in a wider range of genome sequences, we, instead, focused on neighboring genome sequences (the 10 bases upstream and downstream) of all the CpG sites, which we hereafter refer to as "CpG site blocks" (see Methods; Supplemental Table S3). Overlapping CpG site blocks were merged, and the combined length of all the CpG site blocks was 507 Mbp (~18%) in the human genome and ~188 Mbp in the medaka genome, accounting for ~37% of the ~500 Mbp that were aligned between the two medaka genomes. We then noticed that one-third of CpG site blocks were covered by <5 reads and were likely to lack accuracy. Thus, we imposed the stringent requirement that each CpG site had a read coverage of $\geq$5 (see Methods). CpG site blocks that met this condition accounted for ~323 Mbp (~11%) in the human genome and ~139 Mbp (~28%) in the medaka genome, which was sufficient for estimating features of the entire genomes (Supplemental Table S3).

We compared the incidence of genetic variations in hypomethylated regions (methylation level $\leq$ 0.2) with that in hypermethylated regions (methylation level $\geq$ 0.8). Because single-nucleotide mutations by deamination of methylated cytosines in CpG dinucleotides are dominant, to examine single-nucleotide substitution rates in this analysis, we excluded SNPs in CpG dinucleotides and SNPs in TpG/CpA dinucleotides which might make TpG/CpA be CpG. Figure 1A shows that, in human sperm, the incidence rate in hypomethylated regions in the entire genome, 0.054%, was significantly lower than that in hypermethylated ones, 0.081% ($P < 10^{-566}$ by a two-proportion *z*-test) (Supplemental Table S4). This significant increase of incidence rate in hypermethylated regions was also detected in intergenic regions ($P < 10^{-305}$), in exons ($P < 10^{-29}$), and in introns ($P < 10^{-151}$) (Supplemental Table S4). We also observed a similar tendency in the medaka system. For example, Figure 1B illustrates a pair of the homologous regions of the human and medaka genomes where the *RPS13* gene is encoded. We saw four SNPs in the hypermethylated human region and a higher SNP rate in the hypermethylated medaka region. We will check this characteristic precisely in the entire medaka genome in what follows.

In medaka, the methylation state of each CpG site is not necessarily conserved between the two strains (Fig. 1C,D; Supplemental Table S3A), which is supported by 4.46×10^{-4} methylation polymorphisms per CG site per generation in *Arabidopsis thaliana* (Becker et al. 2011; Schmitz et al. 2011). In general, however, hypomethylated and hypermethylated regions were highly conserved between the two inbred strains (Fig. 1C,D) and were also highly correlated ($R^2 = 0.73$) between testes and blastulae (Fig. 1E). Thus, hereafter, we used hypo- and hypermethylated regions that were conserved between the two strains as the representative regions. Analysis of SNPs between Hd-rR and HNI (Sasaki et al. 2009) around CpG blocks revealed that the SNP rate of 1.81% in hypomethylated regions was significantly lower than the 2.78% rate in hypermethylated regions ($P < 10^{-2170}$ by a two-proportion *z*-test) (Fig. 1F; Supplemental Table S4). In addition to the entire genome, these characteristics were also significant in intergenic regions ($P < 10^{-2170}$), in introns ($P < 10^{-589}$), and in exons ($P < 10^{-113}$) in spite of inherently strong purifying selection in exons (Fig. 1F; Supplemental Table S4). Furthermore, genome-wide regions that were differentially methylated in the two strains exhibited a SNP rate of 3.41%, which is significantly higher than the rate of 2.78% in hypermethylated regions ($P < 10^{-442}$) (Supplemental Table S4). This indicates that genomic regions with a relaxed selective pressure display a wider divergence in cytosine methylation.

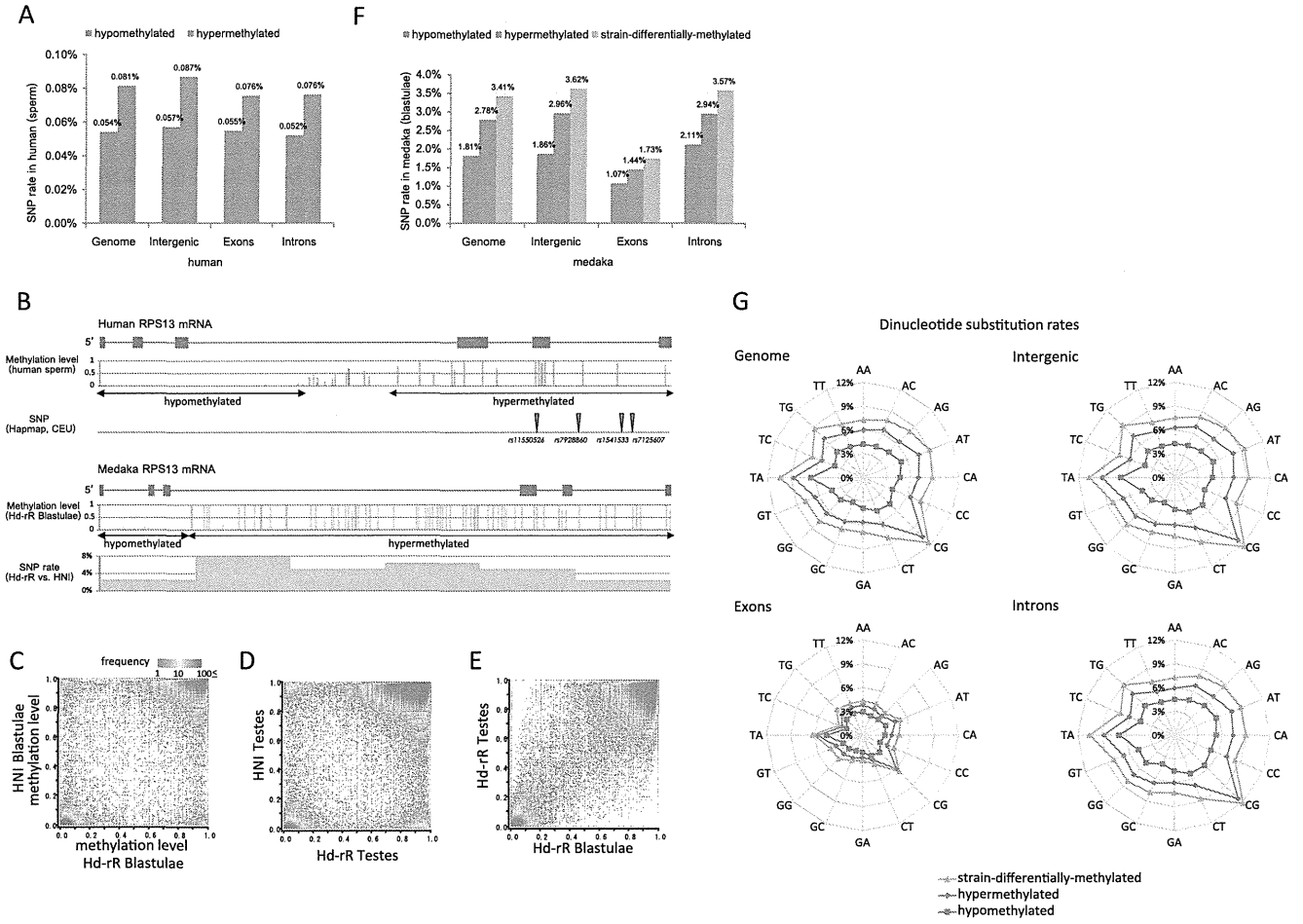

Figure 1. Methylation patterns and substitution rates in the inbred medaka strains, Hd-rR and HNI. (*A*) SNP (single-nucleotide polymorphism) rates in hyper- and hypomethylated CpG blocks in the reference human genome (hg19). The difference in SNP rates was significant in the entire genome ($P < 10^{-566}$ by a two-proportion *z*-test) (Supplemental Table S4), in intergenic regions ($P < 10^{-305}$), in exons ($P < 10^{-29}$), and in introns ($P < 10^{-151}$). (*B*) Methylation level and SNP distribution in the homologous regions of the human and medaka genomes where gene *RPS13* is coded. (*C*,*D*) Comparisons of the methylation patterns in Hd-rR and HNI. The vertical and horizontal axes indicate methylation level. The heat map uses logarithmic coordinates and presents the number of corresponding CpG site blocks. Conserved hypermethylated and hypomethylated patterns between the two strains were dominant, except for a small number of hot spots observed in the differentially methylated regions (differences in methylation level ≥ 0.5). (*E*) Comparison of the methylation patterns in blastulae and testes in Hd-rR. (*F*) SNP rates in hypo-, hyper-, and strain-differentially methylated regions in medaka blastulae grouped by the entire genome, intergenic regions, exons, and introns. The differences between SNP rates of hypo- and hypermethylated regions were remarkable: $P < 10^{-2170}$ (genome), $P < 10^{-2170}$ (intergenic regions), $P < 10^{-113}$ (exons), and $P < 10^{-589}$ (introns) according to a two-proportion *z*-test (Supplemental Table S4). Furthermore, the differences between SNP rates of strain-differentially and hypermethylated regions were also significant (Supplemental Table S4). (*G*) Dinucleotide substitution rates in the whole medaka genome, intergenic regions, exons, and introns in CpG site blocks with various methylation states. Color key presents mutation rates: blue for hypermethylated (methylation level ≥ 0.8 in both strains); red for hypomethylated (methylation level ≤ 0.2 in both strains); and green for strain-differentially methylated (difference in methylation level between the two strains ≥ 0.5) in blastulae. The axes in each radar chart represent substitution rates of individual dinucleotides. Each dinucleotide shows the same substitution rate as its reverse complementary dinucleotide. Significant differences between substitution rates in hypo- and hypermethylated regions were observed for all dinucleotides, and the *P*-values, according to a two-proportion *z*-test, were $P < 10^{-441}$ (genome), $P < 10^{-263}$ (intergenic regions), $P < 10^{-15}$ (exons), and $P < 10^{-69}$ (introns) (Supplemental Table S5).

The high divergence between the two medaka strains provided a detailed analysis of mutations of dinucleotides. Figure 1G shows that the peak substitution rates occurred in CG and TA dinucleotides. The CG substitutions (~11% in hypermethylated regions) were predominantly transitions, which can be explained by the deamination of methylated cytosines that dominates point substitutions in vertebrates (Lindahl and Nyberg 1972; Cooper and Krawczak 1989), whereas in hypomethylated regions, the CG substitution rate was moderate. Another commonly mutated dinucleotide, TA, is universally underrepresented because of its high mutation rate (Ohno 1988; Burge et al. 1992). Our novel finding is that substitution rates for all dinucleotides in CpG site blocks change simultaneously when neighboring CG dinucleotides are hypermethylated, hypomethylated, or differentially methylated. These characteristics were significant for any dinucleotide in the entire genome ($P < 10^{-441}$ by a two proportion *z*-test), in intergenic regions ($P < 10^{-263}$), in introns ($P < 10^{-69}$), and in exons ($P < 10^{-15}$), implying that methylation patterns have a comprehensive impact on genetic variation (see the *P*-values for all dinucleotides in Supplemental Table S5).

One might be concerned that the high CG substitution rate could affect the substitution rates for dinucleotides other than CG;

however, the effects are limited to dinucleotides immediately before CG dinucleotides and are considerably small due to the rarity of CG dinucleotides (2.0% in the medaka genome). In fact, the substitution rates of dinucleotides excluding the dinucleotides overlapping CGs by one base remain essentially unchanged (Supplemental Fig. S3). Another possible explanation for the slower mutation rate in evolutionarily conserved hypomethylated regions is that these regions are around transcription start site (TSS) regions that are highly conserved (Taylor et al. 2006; Sasaki et al. 2009); however, evolution in such regions remained approximately the same even after hypomethylated regions near TSSs (those lying in the 500 bp immediately upstream of and downstream from the TSSs for 23,531 predicted medaka genes) were excluded from the analysis (Supplemental Fig. S4; Supplemental Table S3A,B). The higher substitution rate in evolutionarily conserved hypermethylated regions may be influenced by transposable elements which are highly mutated and methylated in vertebrates (Goll and Bestor 2005). However, transposons in the medaka genome are extremely rare and can be ignored (Kasahara et al. 2007). Therefore, methylation-dependent factors other than the influence of gene transcription or transposons may contribute to the slow evolution of hypomethylated regions.

A related problem is that methylation patterns may be affected by *cis*-acting polymorphisms, such as mutations in CpG sites (Schilling et al. 2009; Shoemaker et al. 2010), but little is known about fundamental *cis*-elements. It is highly nontrivial to identify *cis*-elements that induce the CpG methylation state. We attempted to find sequence motifs that were enriched and evolutionarily conserved in hypomethylated regions because such motifs could be candidate *cis*-elements and would be informative for further studies. We, therefore, searched for potential conserved motifs that were capable of distinguishing two collections of 10,000 hypo- or hypermethylated CpG site blocks using AdaBoost, an established machine learning algorithm (see Methods). The most significant motif was "CGCG," and the incidence of "CGCG" in 10,000 hypomethylated regions was 61.39%, compared with 21.84% in 10,000 hypermethylated regions, showing the significant relevance of "CGCG" sequence enrichment to the methylation state ($P < 10^{-680}$ by a two-proportion *z*-test), as illustrated in Figure 2A. We then examined whether the CGCG motif is evolutionarily conserved in hypomethylated regions. For this purpose, it was necessary to predict the ancestor of Hd-rR and HNI strains. Therefore, we assembled a draft genome of an outgroup medaka strain, HSOK. Figure 2B shows that CGCG is significantly more conserved in hypomethylated regions than in hypermethylated regions ($P = \sim 0.203\%$ by a two-proportion *z*-test). This evolutionarily preserved CGCG motif in hypomethylated regions may be attributable to slow CpG deamination or fast CpG-gaining. To clarify this issue, we calculated dinucleotide-gaining substitution rates by comparing the genomes of Hd-rR and HNI to their ancestor which was estimated from the outgroup HSOK genome (Fig. 2C; see Methods). The rate of CpG gain in hypomethylated regions, 1.4%, was much smaller than the 3.0% rate observed in hypermethylated ones, supporting the hypothesis that slow CpG deamination is the dominant factor for maintaining CGCG richness in hypomethylated regions (Cohen et al. 2011). Meanwhile, we observed an intriguing property in differentially methylated regions in different strains: The rates of CG/CC gain in

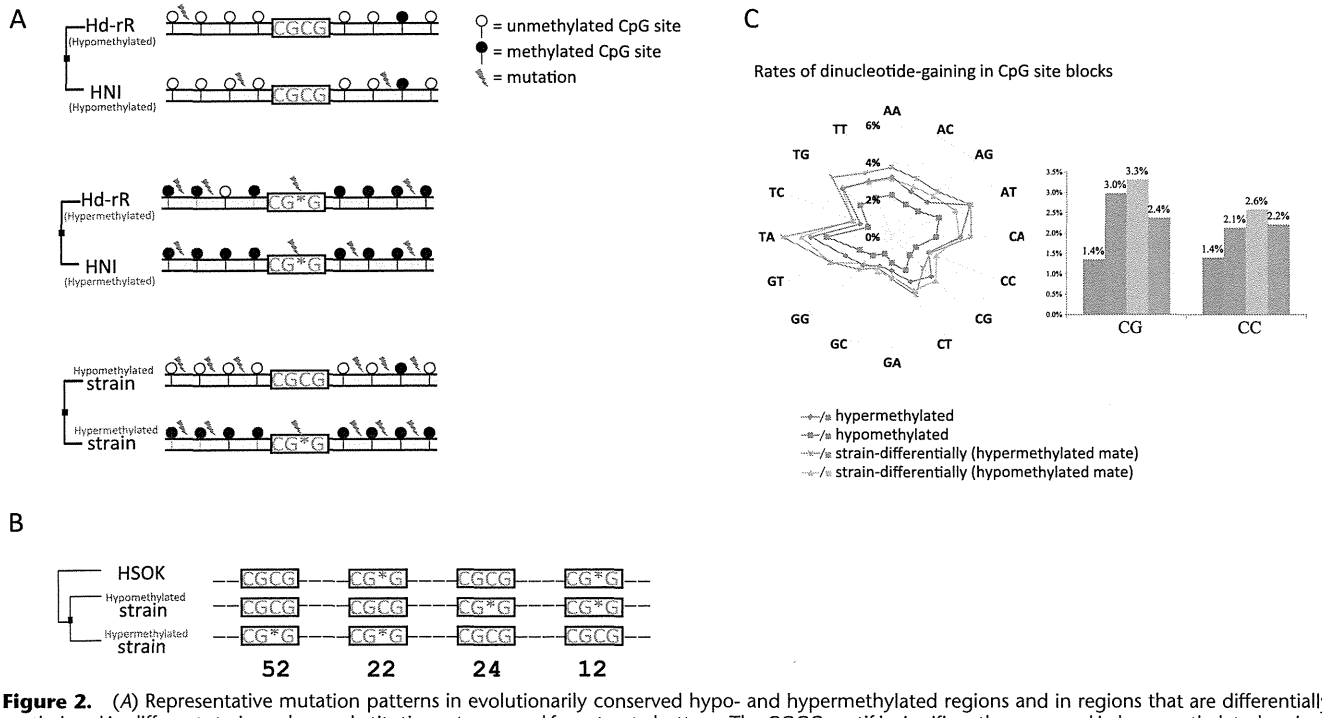

Figure 2. (*A*) Representative mutation patterns in evolutionarily conserved hypo- and hypermethylated regions and in regions that are differentially methylated in different strains, where substitution rates ascend from top to bottom. The CGCG motif is significantly conserved in hypomethylated regions ($P < 10^{-680}$ by a two-proportion *z*-test). (*B*) Number of strain-differentially methylated CpG site blocks with either gain or loss of the CGCG motif. Of 1656 CGCG motif occurrences in multiple alignments of the genomes of the three strains, 52 (24, respectively) were conserved in hypomethylated (hypermethylated) regions of one of Hd-rR or HNI but were mutated in hypermethylated (hypomethylated) regions of the other strain. We then evaluated the significance of the difference between the means in the two groups, 52/1656 vs. 24/1656, to obtain a *P*-value of 0.203% according to a two-proportion *z*-test. (*C*) The rates of dinucleotide gain in CpG site blocks. Gain rates in the hypo- or hypermethylated CpG site blocks and each mate in strain-differentially methylated blocks are shown.

the hypomethylated mate are higher than those in the hypermethylated mate (3.3% and 2.4% compared with 2.6% and 2.2%, respectively), suggesting that CG/CC gains are involved in the evolution of methylation among different strains.

Next, we examined the methylation and genetic variations of two germline-like tissue types (blastulae and testes) and somatic liver cells that are fully differentiated and mostly homogeneous. In this analysis, we focused on tissue types in a single strain, Hd-rR. Figure 3A shows an abundance of CpG site blocks with tissue-differentially methylation states such that differences in methylation levels between each of the germline-like tissues and liver are more than 0.5. Notably, ~99% of tissue-differentially methylated regions were hypomethylated in the liver but were hypermethylated in blastulae and testes. Remarkably, as illustrated in Figure 3B, the tissue-differentially methylated regions showed a significantly lower SNP rate than did the hypomethylated ones ($P < 10^{-82}$) (Supplemental Table S6). Further analysis on dinucleotides revealed that the substitution rates of dinucleotides except for CC/GG/CG in differentially methylated regions were also significantly lower than were those in hypomethylated regions ($P < 10^{-3}$) (Fig. 3C; Supplemental Table S7A), which was also seen even after excluding regions near TSSs (Supplemental Fig. S6). The high CG substitution rate in these regions could have been a consequence of deamination mutations occurring in primary germline-like cells. These results suggest that, because change in DNA methylation can affect transcription so that it can regulate reprogramming during somatic development, a positive selective pressure of genomic sequence could be involved in these differentially methylated regions (Fig. 3D).

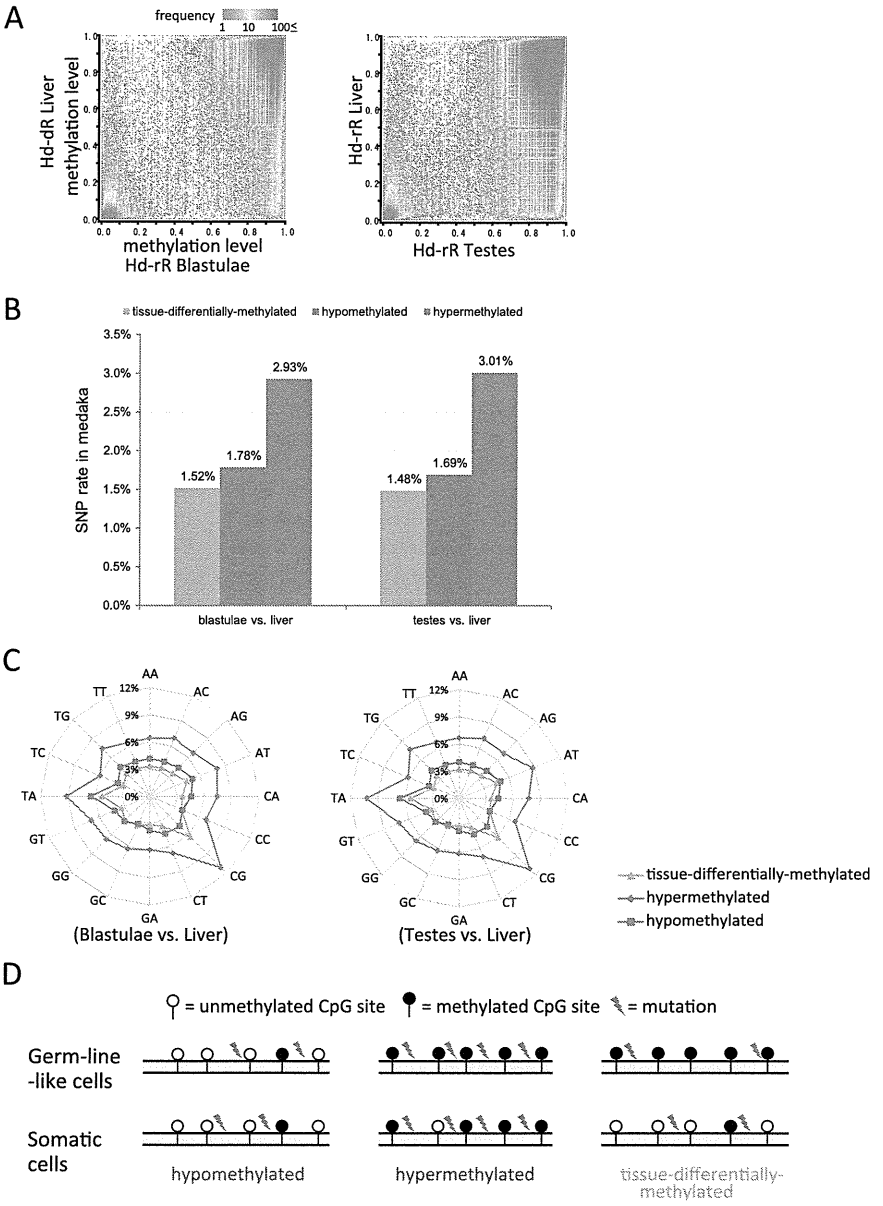

Figure 3. Methylation patterns and substitution rates in different tissue types. (*A*) Comparison of the methylation patterns in Hd-rR: blastulae vs. liver, and testes vs. liver. The vertical and horizontal axes show methylation level. The methylation patterns in HNI are presented in Supplemental Fig. S5 and are similar to those in these figures. (*B*) SNP rates in CpG site blocks with three methylation states: hypomethylated in both of the two tissue types, hypermethylated in both, and differentially methylated between the two tissue types. Significant differences in SNP rates were seen between tissue-differentially and hypomethylated regions ($P < 10^{-82}$ by a two-proportion *z*-test) (Supplemental Table S6), and between hypo- and hypermethylated CpG site blocks ($P < 10^{-2170}$) (Supplemental Table S6). (*C*) Dinucleotide substitution rates in CpG site blocks with the three methylation states. The substitution rates of all dinucleotides, except for CC/GG/CG, in tissue-differentially methylated regions were significantly lower than those in hypomethylated regions ($P < 10^{-3}$ by a two-proportion *z*-test) (Supplemental Table S7A). (*D*) Representative mutation patterns in hypomethylated, hypermethylated, and somatic cell-specific hypomethylated (germline-like-specific hypermethylated) regions. Somatic cell-specific hypomethylated regions exhibited the lowest mutation rates.

Discussion

Because DNA methylation may influence genetic variation, we measured the incidence of genetic variations in methylation states in the human genome as well as in the two medaka inbred strains. We focused on germline-like tissue types, blastulae and testes, to determine the primary effects of methylation and evaluated a total of six single-base-resolution DNA methylomes in the medaka system. Our findings provide novel insights into the relationships between the dynamics of methylation states and genomic variations (Figs. 2A, 3D). Figure 4 illustrates our hypothesis for explaining the difference observed in genetic evolution between hyper- and hypomethylated region. Given that deaminated cytosine (U:G mismatch) is recognized and repaired by base excision repair (BER), deaminated 5-methyl-cytosine (T:G mismatch) would be corrected via more complicated mismatch-repair pathways (Molaro et al. 2011; Fig. 4). Previous studies revealed that two glycosylases, thymine DNA glycosylase (TDG) and methyl-CpG-binding domain

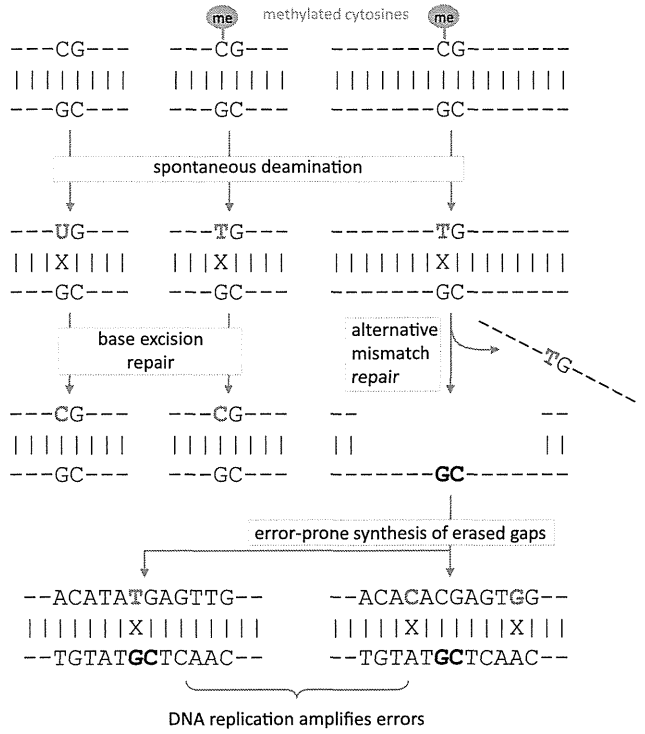

Figure 4. A working model for a higher mutation rate in hypermethylated regions. Deaminated cytosine (U:G mismatch) is repaired by base excision repair (BER), but deaminated 5-methyl-cytosine (T:G mismatch) is corrected by more complicated repair pathways. An alternative mismatch repair system (MMR) might involve low-fidelity DNA polymerase, resulting in the error-prone synthesis of erased gaps.

protein 4 (MBD4), are associated with T:G repair, but processing by these enzymes is biased in the genome and is relatively ineffective (Neddermann et al. 1996; Yoon et al. 2003). An alternative mismatch-repair system (MMR) might gain access to T:G mismatches, which could involve DNA polymerases with lower fidelity (Loeb and Monnat 2008). Thus, we speculate that the error-prone synthesis of the erased gaps around deaminated 5-methyl-cytosines could lead to a higher mutation rate in hypermethylated regions (Fig. 4).

In an effort to search for candidate *cis*-elements for affecting the CpG methylation state, we observed a significant correlation between the methylation state of CpG sites and the presence of CGCG in the neighborhood. This CGCG motif was evolutionarily conserved in hypomethylated regions with statistical significance. A comparative genomic analysis showed slow CpG-gaining in hypomethylated regions, supporting the hypothesis that the evolutionary conservation of CGCG was brought about by slow CpG deamination (Cohen et al. 2011). It is intriguing to ask whether or not CGCG and other candidate motifs are associated with or affect allele-specific methylation. Breeding the F1 generation of the two medaka inbred strains is a promising approach to this problem, and high-throughput short-read bisulfite sequencing is an inexpensive way of observing DNA methylome, though it suffers from a technical difficulty in distinguishing short reads of ~100 bp from two different alleles. The high incidence of genetic variation (~3.4%) between the two inbred strains would facilitate achieving this classification task because an average of 3.4 SNPs per 100-bp read would make it easier to align a read to its originating allele.

Overall, these genetic consequences of methylation, including the slower evolution rate of the *cis*-element "CGCG" found in hypomethylated regions of the medaka system, should also be explored in other species.

Methods

Construction and sequencing of medaka bisulfite-treated DNA libraries

Genomic DNA from medaka tissue cells (blastulae, testes, and liver from two medaka strains, Hd-rR and HNI) was isolated and sonicated to a desired size range (100–400 bp). The DNA fragments were treated with DNA polymerase to generate blunt ends and were ligated with double-stranded DNA adaptors containing methylated cytosines, which were designed to amplify only those DNA fragments carrying bisulfite-converted adaptor sequences at both ends. Followed by 7–10 cycles of PCR, 250 –450-bp size-fractionated DNA was sequenced using an Illumina GAIIx genome analyzer. A validation experiment was performed according to the same procedure using genomic DNA from *S. cerevisiae* S288C.

Computational and statistical methods

We converted all cytosines in reads and in both the Watson and Crick strands of the reference genome to thymines for primary mapping and used Smith-Waterman alignments between the original sequences of primary best hits. All possible methylation patterns were evaluated, and only uniquely mapped reads were retained for further analyses. The level of methylation of a particular cytosine was estimated by dividing the number of mapped reads reporting a cytosine (C) by the total number of reads reporting a C or T (thymine). The comparison of substitution rate and methylation level was performed among CpG site blocks, which consist of at least one CpG site and its surrounding upstream or downstream 10 bases. Overlapping CpG site blocks were merged. A two-proportion *z*-test was performed to test the significance of the substitution rate difference observed between hypo-/hypermethylated CpG site blocks. The AdaBoost algorithm was used to find conserved motifs in hypo-/hypermethylated CpG site blocks. Estimation of the ancestor sequence of Hd-rR and HNI was given by multiple alignments using an out-group medaka strain HSOK which was assembled using the SOAPdenovo assembler (Li et al. 2010). Details of methods can be found in Supplemental Methods.

Data access

All sequence data are deposited at the NCBI Sequence Read Archive (SRA) (http://www.ncbi.nlm.nih.gov/sra) (accession number SRA026693).

Acknowledgments

This work was supported in part by a Grant-in-Aid for Scientific Research (A23241058) from the MEXT, by the Global COE program (Deciphering Biosphere from Genome Big Bang) from the MEXT, and by the Innovative Cell Biology by Innovative Technology from the MEXT. Computational time was provided by the Information Technology Center, the University of Tokyo.

References

Becker C, Hagmann J, Muller J, Koenig D, Stegle O, Borgwardt K, Weigel D. 2011. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480:** 245–249.

Burge C, Campbell AM, Karlin S. 1992. Over-representation and under-representation of short oligonucleotides in DNA-sequences. *Proc Natl Acad Sci* **89:** 1358–1362.
Chandler LA, Ghazi H, Jones PA, Boukamp P, Fusenig NE. 1987. Allele-specific methylation of the human c-Ha-*ras*-1 gene. *Cell* **50:** 711–717.
Cohen NM, Kenigsberg E, Tanay A. 2011. Primate CpG islands are maintained by heterogeneous evolutionary regimes involving minimal selection. *Cell* **145:** 773–786.
Cokus SJ, Feng SH, Zhang XY, Chen ZG, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452:** 215–219.
Cooper DN, Krawczak M. 1989. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Hum Genet* **83:** 181–188.
Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. 1978. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* **274:** 775–780.
Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J, et al. 2009. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat Biotechnol* **27:** 353–360.
Goll MG, Bestor TH. 2005. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* **74:** 481–514.
Heijmans BT, Kremer D, Tobi EW, Boomsma DI, Slagboom PE. 2007. Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human *IGF2/H19* locus. *Hum Mol Genet* **16:** 547–554.
Hellman A, Chess A. 2010. Extensive sequence-influenced DNA methylation polymorphism in the human genome. *Epigenetics Chromatin* **3:** 11. doi: 10.1186/1756-8935-3-11.
Hong Y, Winkler C, Schartl M. 1998. Production of medakafish chimeras from a stable embryonic stem cell line. *Proc Natl Acad Sci* **95:** 3679–3684.
The International HapMap Consortium. 2003. The International HapMap Project. *Nature* **426:** 789–796.
Kasahara M, Naruse K, Sasaki S, Nakatani Y, Qu W, Ahsan B, Yamada T, Nagayasu Y, Doi K, Kasai Y, et al. 2007. The medaka draft genome and insights into vertebrate genome evolution. *Nature* **447:** 714–719.
Kerkel K, Spadola A, Yuan E, Kosek J, Jiang L, Hod E, Li K, Murty VV, Schupf N, Vilain E, et al. 2008. Genomic surveys by methylation-sensitive SNP analysis identify sequence-dependent allele-specific DNA methylation. *Nat Genet* **40:** 904–908.
Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.
Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, et al. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* **20:** 265–272.
Lindahl T, Nyberg B. 1972. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* **11:** 3610–3618.
Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. 2008. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133:** 523–536.
Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462:** 315–322.
Loeb LA, Monnat RJ Jr. 2008. DNA polymerases and human disease. *Nat Rev Genet* **9:** 594–604.
Meaburn EL, Schalkwyk LC, Mill J. 2010. Allele-specific methylation in the human genome: Implications for genetic studies of complex disease. *Epigenetics* **5:** 578–582.
Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD. 2011. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. *Cell* **146:** 1029–1041.
Neddermann P, Gallinari P, Lettieri T, Schmid D, Truong O, Hsuan JJ, Wiebauer K, Jiricny J. 1996. Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J Biol Chem* **271:** 12767–12774.
Ohno S. 1988. Universal rule for coding sequence construction: TA/CG deficiency-TG/CT excess. *Proc Natl Acad Sci* **85:** 9630–9634.
Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2010. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* **327:** 92–94.
Sasaki S, Mello CC, Shimada A, Nakatani Y, Hashimoto S, Ogawa M, Matsushima K, Gu SG, Kasahara M, Ahsan B, et al. 2009. Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science* **323:** 401–404.
Saxonov S, Berg P, Brutlag DL. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci* **103:** 1412–1417.
Schalkwyk LC, Meaburn EL, Smith R, Dempster EL, Jeffries AR, Davies MN, Plomin R, Mill J. 2010. Allelic skewing of DNA methylation is widespread across the genome. *Am J Hum Genet* **86:** 196–212.
Schilling E, El Chartouni C, Rehli M. 2009. Allele-specific DNA methylation in mouse strains is mainly determined by *cis*-acting sequences. *Genome Res* **19:** 2028–2035.
Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, Schork NJ, Ecker JR. 2011. Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334:** 369–373.
Setiamarga DH, Miya M, Yamanoue Y, Azuma Y, Inoue JG, Ishiguro NB, Mabuchi K, Nishida M. 2009. Divergence time of the two regional medaka populations in Japan as a new time scale for comparative genomics of vertebrates. *Biol Lett* **5:** 812–816.
Shoemaker R, Deng J, Wang W, Zhang K. 2010. Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. *Genome Res* **20:** 883–889.
Sved J, Bird A. 1990. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc Natl Acad Sci* **87:** 4692–4696.
Taylor MS, Kai C, Kawai J, Carninci P, Hayashizaki Y, Semple CA. 2006. Heterotachy in mammalian promoter evolution. *PLoS Genet* **2:** e30. doi: 10.1371/journal.pgen.0020030.
Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al. 2001. The sequence of the human genome. *Science* **291:** 1304–1351.
Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39:** 457–466.
Yamada Y, Watanabe H, Miura F, Soejima H, Uchiyama M, Iwasaka T, Mukai T, Sakaki Y, Ito T. 2004. A comprehensive analysis of allelic methylation status of CpG islands on human chromosome 21q. *Genome Res* **14:** 247–266.
Yoon JH, Iwai S, O'Connor TR, Pfeifer GP. 2003. Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. *Nucleic Acids Res* **31:** 5399–5404.
Zhang Y, Rohde C, Tierling S, Jurkowski TP, Bock C, Santacruz D, Ragozin S, Reinhardt R, Groth M, Walter J, et al. 2009. DNA methylation analysis of chromosome 21 gene promoters at single base pair and single allele resolution. *PLoS Genet* **5:** e1000438. doi: 10.1371/journal.pgen.1000438.

Received March 6, 2012; accepted in revised form May 24, 2012.

Immunization with a Recombinant Vaccinia Virus That Encodes Nonstructural Proteins of the Hepatitis C Virus Suppresses Viral Protein Levels in Mouse Liver

Satoshi Sekiguchi[1], Kiminori Kimura[2], Tomoko Chiyo[1], Takahiro Ohtsuki[1], Yoshimi Tobita[1], Yuko Tokunaga[1], Fumihiko Yasui[1], Kyoko Tsukiyama-Kohara[3], Takaji Wakita[4], Toshiyuki Tanaka[5], Masayuki Miyasaka[6], Kyosuke Mizuno[7], Yukiko Hayashi[8], Tsunekazu Hishima[8], Kouji Matsushima[9], Michinori Kohara[1]*

1 Department of Microbiology and Cell Biology, Tokyo Metropolitan Institute of Medical Science, Setagaya-ku, Tokyo, Japan, **2** Division of Hepatology, Tokyo Metropolitan Komagome Hospital, Bunkyo-ku, Tokyo, Japan, **3** Transboundary Animal Diseases Center, Joint Faculty of Veterinary Medicine, Kagoshima University, Korimoto, Kagoshima, Japan, **4** Department of Virology II, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo, Japan, **5** Laboratory of Immunobiology, Department of Pharmacy, School of Pharmacy, Hyogo University of Health Sciences, Chuo-ku, Kobe, Japan, **6** Laboratory of Immunodynamics, Department of Microbiology and Immunology, Osaka University Graduate School of Medicine, Suita, Osaka, Japan, **7** Chemo-Sero-Therapeutic Research Institute, Okubo, Kumamoto, Japan, **8** Department of Pathology, Tokyo Metropolitan Komagome Hospital, Bunkyo-ku, Tokyo, Japan, **9** Department of Molecular Preventive Medicine, School of Medicine, University of Tokyo, Bunkyo-ku, Tokyo, Japan

Abstract

Chronic hepatitis C, which is caused by infection with the hepatitis C virus (HCV), is a global health problem. Using a mouse model of hepatitis C, we examined the therapeutic effects of a recombinant vaccinia virus (rVV) that encodes an HCV protein. We generated immunocompetent mice that each expressed multiple HCV proteins via a Cre/*loxP* switching system and established several distinct attenuated rVV strains. The HCV core protein was expressed consistently in the liver after polyinosinic acid–polycytidylic acid injection, and these mice showed chronic hepatitis C-related pathological findings (hepatocyte abnormalities, accumulation of glycogen, steatosis), liver fibrosis, and hepatocellular carcinoma. Immunization with one rVV strain (rVV-N25), which encoded nonstructural HCV proteins, suppressed serum inflammatory cytokine levels and alleviated the symptoms of pathological chronic hepatitis C within 7 days after injection. Furthermore, HCV protein levels in liver tissue also decreased in a CD4 and CD8 T-cell-dependent manner. Consistent with these results, we showed that rVV-N25 immunization induced a robust CD8 T-cell immune response that was specific to the HCV nonstructural protein 2. We also demonstrated that the onset of chronic hepatitis in CN2-29$^{(+/-)}$/MxCre$^{(+/-)}$ mice was mainly attributable to inflammatory cytokines, (tumor necrosis factor) TNF-α and (interleukin) IL-6. Thus, our generated mice model should be useful for further investigation of the immunological processes associated with persistent expression of HCV proteins because these mice had not developed immune tolerance to the HCV antigen. In addition, we propose that rVV-N25 could be developed as an effective therapeutic vaccine.

Citation: Sekiguchi S, Kimura K, Chiyo T, Ohtsuki T, Tobita Y, et al. (2012) Immunization with a Recombinant Vaccinia Virus That Encodes Nonstructural Proteins of the Hepatitis C Virus Suppresses Viral Protein Levels in Mouse Liver. PLoS ONE 7(12): e51656. doi:10.1371/journal.pone.0051656

Editor: Naglaa H. Shoukry, University of Montreal, Canada

Received March 13, 2012; **Accepted** November 5, 2012; **Published** December 17, 2012

Funding: This study was supported by grants from the Ministry of Education, Culture, Sports, Science, and Technology of Japan; the Program for Promotion of Fundamental Studies in Health Sciences of the Pharmaceuticals and Medical Devices Agency of Japan; and the Ministry of Health, Labor, and Welfare of Japan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: kohara-mc@igakuken.or.jp

Introduction

Hepatitis C virus (HCV) is a major public health problem; approximately 170 million people are infected with HCV worldwide [1]. HCV causes persistent infections that can lead to chronic liver diseases such as chronic hepatitis, liver cirrhosis, and hepatocellular carcinoma (HCC) [2]. Antiviral drugs are not highly effective in individuals with a chronic infection; furthermore, an effective vaccine against HVC has not been developed. A convenient animal model of HCV infection will greatly facilitate the development of an effective HCV vaccine.

Transgenic mice that express HCV proteins have been generated to study HCV expression [3,4]; however, in each of these cases, the relevant transgenes is expressed during embryonic development; therefore, the transgenic mice become immunotolerant to the transgenic products, and consequently, the adult mice are not useful for investigations of the pathogenesis of chronic hepatitis C. To address this problem, we developed a system that can drive conditional expression of an HCV transgene; our system involves the Cre/*loxP* system and a recombinant adenovirus capable of expressing Cre recombinase [5,6]. Concerns have been expressed that an adenovirus and transient expression of HCV proteins could induce immune responses [5] and, therefore, obscure any evidence of the effect of the host immune responses on chronic liver pathology. Therefore, here, we used a Cre/*loxP* switching system to generate an immunocompetent mouse model

of HCV protein expression; with this system, we could study the host immune responses against HCV proteins.

Folgori et al. (2006) reported effective vaccination of chimpanzees with an adenoviral vector and plasmid DNA encoding the HCV nonstructural region. This technique protected the liver tissues from acute hepatitis, which results when whole animals are challenged with virus [7]. However, this vaccine has not yet been shown to be effective against chronic HCV infection.

Here, we aimed to address how HCV expression causes chronic liver diseases and to provide new options for HCV vaccine development. Using LC16m8, a highly attenuated strain of vaccinia virus (VV), we generated three recombinant vaccinia viruses (rVVs) that each encoded one of three different HCV proteins and found that one recombinant virus (rVV-N25), which encoded nonstructural HCV proteins, resolved pathological chronic hepatitis C symptoms in the liver. We also found that immunization with rVV-N25 suppressed HCV core protein levels in the livers of transgenic mice; moreover, this suppression was mediated by CD4 and CD8 T cells, as has been previously reported [8].

Results

Generation of a Model of Persistent HCV Protein Expression

To produce adult mice that express an HCV transgene, we bred CN2-29 transgenic mice, which carry an HCV transgene, [5,6,9] with Mx1-Cre transgenic mice [10], which express Cre recombinase in response to interferon (IFN)-α or a chemical inducer of IFN-α, poly(I:C) (Figure 1A). Following poly(I:C) injection, the HCV transgene was rearranged, and HCV sequences were expressed in the livers of F1 progeny (CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice) within 7 days after poly(I:C) injection (Figure 1B).

To evaluate the characteristic features of these CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice, we analyzed serum alanine aminotransferase (ALT) and liver HCV core protein levels after poly(I:C) injection. As illustrated in Figure 1C, serum ALT levels increased and reached a peak at 24 h after the first poly(I:C) injection; this elevation appeared to be a direct result of the poly(I:C) treatment, which causes liver injury [11]. After this peak, serum ALT levels dropped continuously until day 4, and then ALT levels began to increase, as did HCV core protein levels. Thereafter, the HCV core protein was expressed consistently for at least 600 days.

Histological analysis showed HCV core protein expression in most hepatocytes of the transgenic mice; these mice showed evidence of lymphocytic infiltration that was caused by the HCV core proteins (Figure 1D and E). These observations, in addition to the modified histology activity index (HAI) scores, indicated that expression of HCV proteins caused chronic hepatitis in the CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice because a weak, though persistent, immune response followed an initial bout of acute hepatitis (Figure S1). Moreover, we observed a number of other pathological changes in these mice – including swelling of hepatocytes, abnormal architecture of liver-cell cords, abnormal accumulation of glycogen, steatosis, fibrosis, and HCC (Figures 1E and F, Table S1). Steatosis was mild in the younger mice (day 21) and became increasingly severe over time (days 120 and 180; Figure S2). Importantly, none of the pathological changes were observed in the CN2-$29^{(+/-)}$/$MxCre^{(-/-)}$ mice after poly(I:C) injection (Figure 1F).

Recombinant Vaccinia Virus Immunization in HCV Transgenic Mice

To determine whether activation of the host immune response caused the reduction with HCV protein levels in the livers of CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice, we used a highly attenuated VV strain, LC16m8, to generate three rVVs [12]. Each rVV encoded a different HCV protein; rVV-CN2 encoded mainly structural proteins, rVV-N25 encoded nonstructural proteins, and rVV-CN5 encoded the entire HCV protein region (Figure 2A). Because rVVs can express a variety of proteins and induce strong and long-term immunity, they have been evaluated as potential prophylactic vaccines [13].

We used western blots to confirm that each HCV protein was expressed in cell lines. Each of seven proteins – the core, E1, E2, NS3-4A, NS4B, NS5A, and NS5B – was recognized and labeled by a separate cognate antibody directed (Figure S3). To induce effective immune responses against HCV proteins in transgenic mice, we injected an rVV-HCV (rVV-CN2, rVV-CN5, or rVV-N25) or LC16m8 (as the control) intradermally into CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice 90 days after poly(I:C) injection (Figure 2B). Analysis of liver sections 7 days after immunization with rVV-N25 revealed dramatic improvement in a variety of pathological findings associated with chronic hepatitis – including piecemeal necrosis, hepatocyte swelling, abnormal architecture of liver-cell cords, abnormal accumulation of glycogen, and steatosis (Figures 2C–E). Collectively, these results demonstrated that only the rVV-N25 treatment resulted in histological changes indicative of improvement in the chronic hepatitis suffered by the transgenic mice.

To determine whether rVV-N25 treatment induced the same effect in other strains of HCV transgenic mice, we analyzed RzCN5-$15^{(+/-)}$/$MxCre^{(+/-)}$ mice, which express all HCV proteins; in these mice, chronic hepatitis was resolved within 28 days of immunization with rVV-N25. Taken together, these findings indicated that rVV-N25 had a dramatic therapeutic effect on both types of HCV transgenic mice (Figure S4).

Treatment with rVV-N25 Reduced the HCV Core Protein Levels in the Livers

To assess in detail the effects of rVV-HCV immunization on HCV protein clearance from the livers of CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice, we monitored the levels of HCV core protein in liver samples via ELISA. We found that within 28 days after immunization the HCV core protein levels were significantly lower in livers of rVV-N25-treated mice than in those of control mice (Figure 3A). Immunohistochemical analysis indicated that, within 28 days after immunization, levels of HCV core protein were substantially lower in the livers of CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice than in those of control mice (Figure 3B). Importantly, neither resolution of chronic hepatitis nor reduction in the HCV protein levels was observed in the mice treated with LC16m8, rVV-CN2, or rVV-CN5. These results indicated that HCV nonstructural proteins might be important for effects of therapeutic vaccines. In contrast, rVV-CN5 which encoded HCV structural and non-structural proteins did not show any significant effects. These results indicated that HCV structural proteins might have inhibited the therapeutic effects of the non-structural proteins. Therefore, it may be important to exclude the HCV structural proteins (aa 1–541) as antigenic proteins when developing therapeutic vaccines against chronic hepatitis C.

In addition, we measured serum ALT levels in CN2-$29^{(+/-)}$/$MxCre^{(+/-)}$ mice from all four treatment groups 28 days after rVV-HCV immunization. Serum ALT levels were not significant-