Step 1: Deletion of a non-significant edge from the constructed network model

Use 0.05 as the significance level for the determination of the significant regulation among the variables. After the parameters are estimated, the inverse matrix of the Fisher information matrix of parameters is calculated. The inverse matrix of Fisher information represents the asymptotic parameters' covariance matrix. The probability of each parameter is calculated by using this asymptotic parameters' matrix, since all of the parameters are usually normally distributed.

Step 2: Reconstruction of the network model

The structure of the network model without the non-significant edge is completely different from that of the former model. Thus, all parameters should be re-calculated from the reconstructed model, and the similarity of the network structure should also be re-calculated.

Step 3: Iteration of Steps 1 and 2 until all edges become significant

Since the probabilities of all of the edges in the reconstructed models have also changed, the deletion of the non-significant edges is executed step-by-step.

Step 4: Addition of a possible causal edge to the reconstructed model

According to the Modification Index (MI), we add a new causal edge between the observed variables. The MI measures how much the chi-square statistic is expected to decrease if a particular parameter setting is constrained (Joreskog & Sorbom, 1984). The MI value indicates the possibility of new causality between the variables, and thus we add a new edge according to the highest MI score.

Step 5: Iteration from Steps 1 to 3

The addition of a new edge to a constructed model changes the structure of the network model again. In other words, all parameters, including the probabilities of all edges, have also changed. Thus, we execute the iteration from Step 1 to Step 3 again.

Step 6: Determination of significant relationships among error terms

After all of the edges are significant and all of the MI scores are lower than 10.0 in the constructed model, the significant relationships between the error terms are estimated by the MI scores. The relationships among the error terms have no direction, and thus they are a correlation between error terms. The relationships between the error terms were considered to be other regulatory systems in the living cell. Thus, these relationships among the error terms were used for the calculations, but were not incorporated into the network, and thus they have been excluded from the figures.

3. Results

3.1 Initial Model Assumption

To construct the initial network model of each chemical, we utilized our newly developed method. One of the distinguishing features of our new method is its ability to include the cyclic structure in the network model. Cyclic regulation, such as feedback regulation, is considered to be important for living cells to control normal gene expression, and the new method is useful to detect the cyclic regulation from the gene expression data. The initially constructed models are shown in Figure 2. The initial model of TCDD was the most complex structure. The components of the constructed models were 9 genes with 19 relationships in Acrylamide, 8 genes with 12 relationships in Diethylnitrosamine, 9 genes with 23 relationships in TCDD, and 8 genes with 10 relationships in Thalidomide.

There are some obvious features in the network diagram of each initial model. The numbers of exogenous and endogenous genes are different from each other. In the initial Acrylamide model, four genes were arranged as exogenous variables, but only Oct3/4 was arranged as the last endogenous variable. Thus, it is considered that acrylamide quickly affected the expression of many genes, and only one gene was affected later. In contrast, only one gene was arranged as an exogenous variable and many genes were arranged as the last endogenous variables in the initial Thalidomide model. These differences between the initial chemical models summarized the distinctive gene expression profiles for each chemical. The initial TCDD model involved some cyclic regulation, even though the other models had only hierarchical regulation.

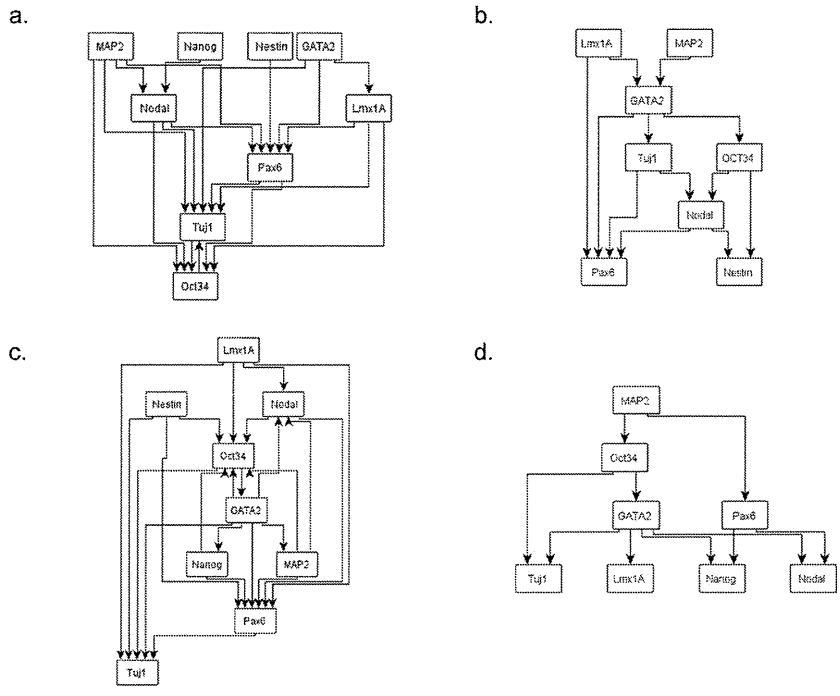

Figure 2. Initial network models

The initial models of the selected chemicals were constructed by the developed approach. One initial model was constructed for each chemical, since the initial model included summarized time-series information and concentration information. (a) Initial model constructed from all gene expression profiles with all Acrylamide exposure. (b) Initial model of Diethylnitrosamine. (c) Initial model of TCDD. (d) Initial model of Thalidomide. The numbers of genes in the initial models were 9 in Acrylamide, 8 in Diethylnitrosamine, 9 TCDD, and 8 in Thalidomide.

Before the calculation of SEM, all of the initial models were simplified, since the initial models included some duplicated interactions among the genes, such as direct interactions between two genes and indirect interactions between them. In the simplification process for the initial models, the longest path between two genes was retained, since the arrows indicated only time precedence, not causalities in the initial model. Therefore, the difference between direct and indirect interactions is not important. By retaining the longest paths, all of the preceding information was included, as the simplest diagram.

3.2 Inferred Networks by SEM

The final inferred networks for each chemical and the goodness of fit scores are depicted in Figure 3, and the estimated regression weights of the edges are displayed in Table 1. The inferred networks of the chemicals revealed distinct structures. The differences between the gene regulation by chemicals were clarified by the shapes of the inferred network models. The Acrylamide network was a centralized model, the Diethylnitrosamine network was a ladder-like model, the TCDD network was a closed circular structure, and the Thalidomide network was a diffusion type.

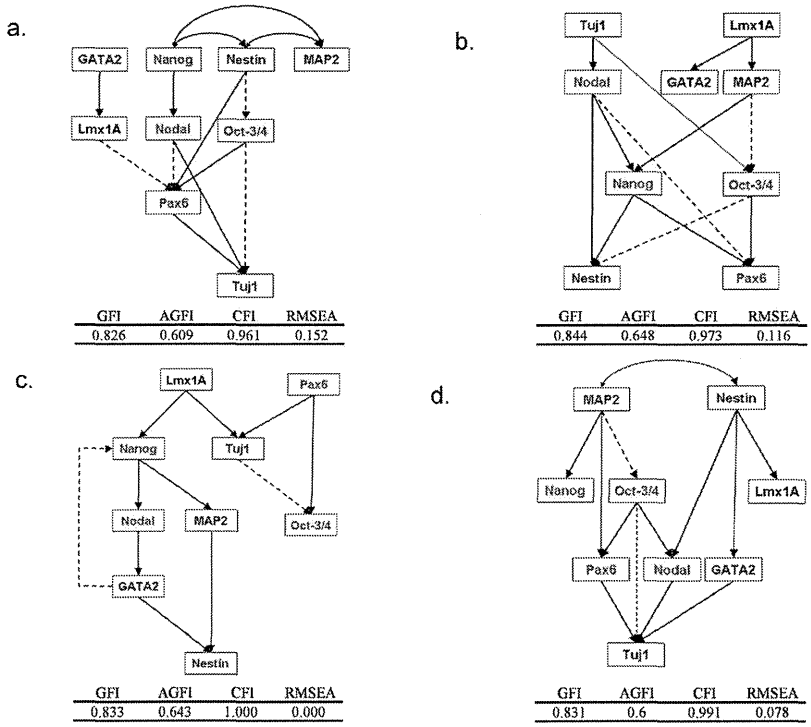

Figure 3. Inferred Toxic-dependent Networks

The optimal model for each chemical, obtained by the developed SEM iteration procedure. A positive relationship between genes is displayed with a solid arrow. A negative relationship between genes is displayed with a dashed arrow. Gene names with blue characters indicate "neurodevelopment related genes", genes with red characters indicate "cell differentiation-related genes" and genes with black characters indicate "related to transcription of insulin". (a) Acrylamide model; (b) Diethylnitrosamine model; (c) TCDD model and (d) Thalidomide model. The fitting scores are displayed under each model.

One of the unique features of the inferred Acrylamide network was that many genes were arranged at the top phase in the regulatory network, and only one gene was arranged as the final result of all regulation in the network. On the other hand, the shape of the Diethylnitrosamine network looked like a ladder, and two serial regulations interacted with each other. One serial regulation started from Lmx1A, and the other started from Tuj1. These top phase genes were considered as signal input genes, and they were different from those in the Acrylamide and Thalidomide networks. For example, Tuj1 was arranged as a signal input gene in the Diethylnitrosamine network, but it was arranged as an output object in the Acrylamide and Thalidomide networks. The unique feature of the TCDD network is the involvement of some closed circular structures in the inferred model. Among the parts of the circular structure, the regulatory direction from GATA2 to Nodal was different from the other relationships. Furthermore, the regression weight between GATA2 and Nodal was estimated as a negative value. Thus, it was considered that the inferred regulation from GATA2 to Nodal reflected feedback control by GATA2. In the Thalidomide network, the shape of the network model was reversed, as compared to that of the Acrylamide network. Only two genes were arranged at the top phase in the regulatory network, but many genes were arranged at the middle phase in the model. This means that only a few genes are directly affected by thalidomide, but finally many genes are affected throughout the gene regulatory network.

Table 1. Regression weight and probability of each edge

Acrylamide				Diethylnitrosamine			
Parent	**Child**	**Regression Weight**	**P**	**Parent**	**Child**	**Regression Weight**	**P**
GATA2	Lmx1A	0.921	***	Tuj1	Nodal	0.702	***
Nanog	Nodal	0.522	0.003	Lmx1A	MAP2	0.378	0.003
Nestin	Oct-34	-0.437	0.01	Tuj1	Oct-34	0.63	***
Nestin	Pax6	0.64	***	MAP2	Oct-34	-0.475	***
Nodal	Pax6	-0.803	***	Nodal	Nanog	0.295	***
Lmx1A	Pax6	-0.232	***	MAP2	Nanog	0.754	***
Oct-34	Pax6	0.592	***	Lmx1A	GATA2	0.636	***
Nodal	Tuj1	0.843	***	Nodal	Nestin	0.33	***
Pax6	Tuj1	1.09	***	Nodal	Pax6	-0.209	0.005
Oct-34	Tuj1	-0.702	***	Nanog	Pax6	0.418	***
				Nanog	Nestin	0.902	***
				Oct-34	Pax6	1.11	***
				Oct-34	Nestin	-0.193	***

TCDD				Thalidomide			
Parent	**Child**	**Regression Weight**	**P**	**Parent**	**Child**	**Regression Weight**	**P**
GATA2	Nanog	-0.787	***	MAP2	Oct-34	-0.443	0.023
GATA2	Nestin	0.22	***	MAP2	Pax6	0.349	0.005
Lmx1A	Nanog	1.374	***	Nestin	Nodal	1.03	***
Lmx1A	Tuj1	0.476	0.004	Nestin	GATA2	0.664	***
MAP2	Nestin	0.906	***	Oct-34	Pax6	0.932	***
Nanog	MAP2	1.024	***	Oct-34	Nodal	0.258	***
Nanog	Nodal	0.967	***	Oct-34	Tuj1	-0.597	***
Nodal	GATA2	0.931	***	Pax6	Tuj1	1.12	***
Pax6	Oct-34	0.988	***	Nodal	Tuj1	0.349	***
Pax6	Tuj1	0.5	0.003	GATA2	Tuj1	0.167	0.015
Tuj1	Oct-34	-0.324	***	MAP2	Nanog	0.84	***
				Nestin	Lmx1A	0.842	***
				Tuj1	Nanog	0.196	0.002

4. Discussion

Our inferred model revealed the differences between the gene regulation by environmental chemicals. Furthermore, the shapes of the network models reflected the different features of the chemical toxicities well. In the Acrylamide network, the effects of acrylamide toxicity finally aggregated to Tuj1, which is known to contribute to microtubule stability in neuronal cells (Rosenstein et al., 2003). Acrylamide is neurotoxic, and thus it is reasonable that its effect finally aggregated to a neuronal cell-related gene.

In the Diethylnitrosamine network, the cell differentiation genes were arranged from the middle to lower steps. This means that diethylnitrosamine disturbed normal cell differentiation in the embryonic stem cell. These harmful effects were considered to be caused by the carcinogenic genotoxicity of diethylnitrosamine (Ito et al., 1992; Puatanachokchai et al., 2006; Iatropoulos et al., 2006). On the other hand, the neuronal-related genes were arranged at a later phase in the TCDD network model. Although both diethylnitrosamine and TCDD have the same carcinogenic toxicities, their regulatory mechanisms were different.

From the Thalidomide network, it was considered that the receptors of thalidomide toxicity may be rarer than those of other chemicals. However, several types of genes are finally affected by thalidomide chemical toxicity. Among the cell differentiation genes, Nodal and Nanog are important for normal early embryonic development. Nodal is related to the development of the left-right axial structure (Hamada et al., 2002; Grandel & Patel, 2009), and its signaling pathway is important very early in development, for cell fate determination and many other developmental processes (Grandel & Patel, 2009). Nanog is a key factor for maintaining pluripotency in embryonic stem cells (Mitsui, 2003; Chambers et al., 2003). According to the abnormal expression of these cell differentiation-related genes, the thalidomide phenotype, with its harmful side effects, may occur in the human embryo.

We applied an improved SEM approach to reconstruct a gene regulatory model from the gene expression data in human embryonic stem cells, and we have shown that SEM is a powerful approach to estimate the gene

regulation caused by chemical toxicity. The inferred networks clarified the differences between the gene regulation by chemicals, and the features of the chemical toxicities were well reflected in the network structures. Thus, the network construction by SEM is one of the useful approaches for inferring the regulatory relationships among genes. Furthermore, the inferred network among genes can be utilized for the estimation of a chemical's effect, from experimentally obtained expression profiles. The ability to identify expression profiles and the corresponding biological functions is expected to provide further possibilities for SEM in the inference of the effects of chemical toxicity on regulatory mechanisms.

Acknowledgements

We would like to express our gratitude to Dr. Yamane (Kyoto Univ.) and Dr. Ohsako (Univ. of Tokyo) for providing the expression profiles in human embryonic stem cells exposed to 15 chemicals. Their support was quite valuable for this investigation.

References

Aburatani, S. (2012). Network Inference of pal-1 Lineage-Specific Regulation in the C. elegans Embryo by Structural Equation Modeling. *Bioinformation, 8*(14), 652-657. http://dx.doi.org/10.6026/97320630008652

Aburatani, S., & Horimoto, K. (2005). Elucidation of the Relationships between LexA-Regulated Genes in the SOS response. *Genome Informatics, 16*(1), 95-105.

Aburatani, S., Kuhara, S., Toh, H., & Horimoto, K., (2003). Deduction of a gene regulatory relationship framework from gene expression data by the application of graphical Gaussian modeling. *Signal Processing, 83*, 777-788. http://dx.doi.org/10.1016/S0165-1684(02)00476-0

Aburatani. S. (2011). Application of structure equation modeling for inferring a serial transcriptional regulation in yeast. *Gene. Regul. Syst. Bio., 5*, 75-88. http://dx.doi.org/10.4137/GRSB.S7569

Akutsu, T., Miyano, S., & Kuhara, S. (2000). Algorithms for identifying Boolean networks and related biological networks based on matrix multiplication and fingerprint function. *J. Comput. Biol., 7*, 331-343. http://dx.doi.org/ 10.1145/332306.332317

Baccarelli, A., & Bollati, V. (2009). Epigenetics and environmental chemicals. *Curr. Opin. Pediatr., 21*(2), 243-251. http://dx.doi.org/10.1097/MOP.0b013e32832925cc

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley-Interscience.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., & Smith, A. (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell, 113*(5), 643-655. http://dx.doi.org/10.1016/S0092-8674(03)00392-1

Duncan, O. D. (1975). *Introduction to Structural Equation Models*. New York: Academic Press.

Friedman, N., Linial, M., Nachman, I., & Pe'er, D. (2000). Using Bayesian networks to analyze expression data. *J. Comput. Biol., 7*, 601-620. http://dx.doi.org/10.1145/332306.332355

Fujibuchi, W. (2011). Prediction of Chemical Toxicity by Network-based SVM on ES-cell Validation System. *The Proc. of the 2011 Joint Conference of CBI-Society and JSBi.*

Grandel, C., & Patel, N. H. (2009). Nodal signaling is involved in left-right asymmetry in snails. *Nature, 457*(7232), 1007-1011. http://dx.doi.org/10.1038/nature07603

Gündel, U., Benndorf, D., Bergen, M., Altenburger, R., & Küster, E. (2007). Vitellogenin cleavage products as indicators for toxic stress in zebra fish embryos: A proteomic approach. *Proteomics, 7*(24), 4541-4554. http://dx.doi.org/10.1002/pmic.200700381

Hamada, H., Meno, C., Watanabe, D., & Saijoh, Y. (2002). Establishment of vertebrate left-right asymmetry. *Nat. Rev. Genet., 3*(2), 103-113. http://dx.doi.org/10.1038/nrg732

Harrill, J. A., Robinette, B. L., & Mundy, W. R. (2011). Use of high content image analysis to detect chemical-induced changes in synaptogenesis in vitro. *Toxicol. In. Vitro., 25*(1), 368-387. http://dx.doi.org/10.1016/j.tiv.2010.10.011

He, X., Imanishi, S., Sone, H., Nagano, R., Qin, X-Y., Yoshinaga, J., ... Ohsako, S. (2012). Effects of methylmercury exposure on neuronal differentiation of mouse and human embryonic stem cells. *Toxicol. Lett., 212*, 1-10. http://dx.doi.org/10.1016/j.toxlet.2012.04.011

Hou, L., Zhang, X., Wang, D., & Baccarelli, A. (2012). Environmental chemical exposures and human epigenetics. *Int. J. Epidemiol., 41*(1), 79-105. http://dx.doi.org/10.1093/ije/dyr154

Iatropoulos, M. J., Wang, C. X., Keutz, K. E., & Williams, G. M. (2006). Assessment of chronic toxicity and carcinogenicity in an accelerated cancer bioassay in rats of Nifurtimox, an antitrypanosomiasis drug. *Exp. Toxicol. Pathol., 57*(5-6), 397-404. http://dx.doi.org/10.1016/j.etp.2006.01.005

Ito, N., Hasegawa, R., Imaida, K., Masui, T., Takahashi, S., & Shirai, T. (1992). Pathological markers for non-genotoxic agent-associated carcinogenesis. *Toxicol. Lett., 64-65*, 613-620. http://dx.doi.org/10.1016/0378-4274(92)90239-G

Joreskog, K. G., & Sorbom, D. (1984). LISREL-VI: Analysis of Linear Structural Relationships By the Method of Maximum Likelihood. Redondo Beach: Doss-Haus Books.

Li, L., & Caldwell, G. E. (1999). Coefficient of cross correlation and the time domain correspondence. *J. of Electromyography and Kinesiology, 9*, 385-389. http://dx.doi.org/10.1016/S1050-6411(99)00012-7

Mitsui, K. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell, 113*(5), 631-642. http://dx.doi.org/10.1016/S0092-8674(03)00393-3

Pearl, J. (2001). *Causality: Models, Reasoning, and Inference* (2nd ed.). Cambridge: Cambridge University Press.

Puatanachokchai, R., Kakuni, M., Wanibuchi, H., Kinoshita, A., Kang, J. S., Salim, E. I., ... Fukushima, S. (2006). Lack of promoting effects of phenobarbital at low dose on diethylnitrosamine-induced hepatocarcinogenesis in TGF-alpha transgenic mice. *Asian Pac. J. Cancer Prev., 7*(2), 274-278.

Rappolee, D. A., Xie, Y., Slater, J. A., Zhou, S., & Puscheck, E. E. (2012). Toxic stress prioritizes and imbalances stem cell differentiation: implications for new biomarkers and in vitro toxicology tests. *Syst. Biol. Reprod. Med., 58*(1), 33-40. http://dx.doi.org/10.3109/19396368.2011.647381

Rosenstein, J. M., Mani, N., Khaibullina, A., & Krum, J. M. (2003). Neurotrophic effects of vascular endothelial growth factor on organotypic cortical explants and primary cortical neurons. *J. Neurosci., 23*(35), 11036-11044.

Spirtes, P., Glymour, C., & Scheines, R. (2001). *Causation, Prediction, and Search* (2nd ed.). Cambridge: The MIT Press.

Tatsuta, N., Nakai, K., Murata, K., Suzuki, K., Iwai-Shimada, M., Yaginuma-Sakurai, K., ... Satoh, H. (2012). Prenatal exposures to environmental chemicals and birth order as risk factors for child behavior problems. *Environ. Res., 114*, 47-52. http://dx.doi.org/10.1016/j.envres.2012.02.001

Thompson, J., & Bannigan, J. (2008). Cadmium: toxic effects on the reproductive system and the embryo. *Reprod. Toxicol., 25*(3), 304-315. http://dx.doi.org/10.1016/j.reprotox.2008.02.001

Yuan, Y. (2012). Methylmercury: a potential environmental risk factor contributing to epileptogenesis. *Neurotoxicology, 33*(1), 119-126. http://dx.doi.org/10.1016/j.neuro.2011.12.014

Carcinogenesis vol.33 no.12 pp.2311–2319, 2012
doi:10.1093/carcin/bgs347
Advance Access publication November 1, 2012

REVIEW

Splice isoforms as therapeutic targets for colorectal cancer

Koh Miura*, Wataru Fujibuchi[1] and Michiaki Unno

Department of Surgery, Tohoku University Graduate School of Medicine, Sendai 980–8574, Japan and [1]Center for iPS Cell Research and Application, Kyoto University, Kyoto 606–8507, Japan

*To whom correspondence should be addressed. Tel: +81 22 717 7205; Fax: +81 22 717 7209;
Email: k-miura@surg1.med.tohoku.ac.jp

Alternative pre-mRNA splicing allows exons of pre-mRNA to be spliced in different arrangements to produce functionally distinct mRNAs. More than 95% of human genes encode splice isoforms, some of which exert antagonistic functions. Recent studies revealed that alterations of the splicing machinery can cause the development of neoplasms, and understanding the splicing machinery is crucial for developing novel therapeutic strategies for malignancies. Colorectal cancer patients need novel strategies not only to enhance the efficacy of the currently available agents but also to utilize newly identified therapeutic targets. This review summarizes the current knowledge about the splice isoforms of *VEGFA*, *UGT1A*, *PXR*, *cyclin D1*, *BIRC5* (*survivin*), *DPD*, *K-RAS*, *SOX9*, *SLC39A14* and other genes, which may be possible therapeutic targets for colorectal cancer. Among them, the *VEGFA* splice isoforms are classified into VEGFAxxx and VEGFAxxxb, which have proangiogenic and antiangiogenic properties, respectively; *UGT1A* is alternatively spliced into UGT1A1 and other isoforms, which are regulated by pregnane X receptor isoforms and undergo further splicing modifications. Recently, the splicing machinery has been extensively investigated and novel discoveries in this research field are being reported at a rapid pace. The information contained in this review also provides suggestions for how therapeutic strategies targeting alternative splicing can be further developed.

Introduction

Colorectal cancer (CRC) is the third most common cancer worldwide and the second most common cause of cancer mortality; ~608 000 deaths are attributed to this disease annually (1). As indicated in the National Comprehensive Cancer Network (NCCN) guidelines version 1.2013 (2), the anticancer agents used for CRC include 5-fluorouracil (5-FU), irinotecan hydrochloride (CPT-11), oxaliplatin and molecularly targeted agents such as bevacizumab, cetuximab and panitumumab. According to the guidelines, in the adjuvant setting for stage II or III patients with resectable CRC, regimens including 5-FU-based agents and leucovorin with or without oxaliplatin are recommended, and for stage IV patients with unresectable CRC, regimens including 5-FU-based agents and leucovorin with oxaliplatin or CPT-11 and the addition of bevacizumab, cetuximab or panitumumab are considered (2). Over the past two decades, there have been advances in the treatment of CRC; however, patients need novel strategies not only to enhance the efficacy of the above agents but also to utilize newly identified therapeutic targets.

Alternative precursor messenger RNA (pre-mRNA) splicing is the process by which the exons of pre-mRNA are spliced in different arrangements to produce structurally and functionally distinct mRNAs and proteins (Figure 1A) (3). After the completion of the Human Genome Project in 2004, alternative splicing has been recognized as one of the most important mechanisms that maintains genomic and functional diversity. It is well known that >95% of human genes encode splice isoforms (4), some of which exert antagonistic functions. A recent study revealed that alterations of the splicing machinery can cause the development of myeloid neoplasms (5), and understanding the splicing machinery is crucial for the development of novel therapeutic strategies for malignancies. Another recent study revealed that a series of functionally associated splice isoforms are simultaneously expressed under a common regulatory network (6), which supports the notion that an entire set of splice isoforms or their common regulatory network should be considered as therapeutic targets, rather than focusing on a single gene as a target. In this review, we summarize the current knowledge about the potential of using splice isoforms as therapeutic targets, focusing on CRC, and discuss the future work that needs to be done to develop therapeutic strategies targeting these splice isoforms.

Alternative pre-mRNA splicing

The first studies on pre-mRNA splicing were published in 1977 (7,8). Two regulatory factors have been the focus of most investigations related to the splicing machinery: cis-elements and trans-elements (Figure 1A). Among the cis-elements, consensus splice sites such as the 5′ splice site (5′SS; also known as a splice donor site), the branch point motif, the poly-pyrimidine tract [(Y)n] and the 3′ splice site (3′SS; also known as a splice acceptor site) are essential for pre-mRNA splicing. Splice enhancers and silencers are also categorized into cis-elements, both of which are important for the recognition of the 5′SS and 3′SS sites. Depending on their localization within the genome, splice enhancers and silencers are subclassified into exonic splice enhancers (ESEs), intronic splice enhancers (ISEs), exonic splice silencers (ESSs) and intronic splice silencers (ISSs). Cis-elements are bound by trans-elements. Among the trans-elements, spliceosomes are multicomponent complexes comprising >200 subunits. Among the subunits of spliceosomes, serine-/arginine-rich proteins (SR proteins, SRp) predominantly bind to ESEs and ISEs; in contrast, heterogeneous nuclear ribonucleoproteins (hnRNPs) commonly bind to ESSs and ISSs. In many cases, hnRNPs block spliceosome assembly, resulting in exon skipping. Recently, tissue- or organ-specific SR proteins and hnRNPs have been extensively investigated (9,10). Figure 1B shows several patterns of alternative splicing in which splice isoforms are generated: (i) exon skipping in which an alternative exon is excluded or included, (ii) intron retention between constitutive exons, (iii) inclusion of one of the exons in a mutually exclusive manner, (iv) use of alternative 5′SSs, (v) alternative 3′SSs, (vi) alternative initiation sites and (vii) alternative polyadenylation sites.

The normal expression profile is indicated in the lower left panel, whereas the aberrant splicing that is observed in malignancies is shown in the lower right panel of Figure 1A, and can be subclassified into two categories: (i) aberrant splice isoforms as individual transcripts and (ii) an aberrant expression profile of splice isoforms as an entire set of transcripts; both of which occur at the germ cell or somatic cell level. Herein the word 'change' is used to encompass both 'genetic polymorphism' and 'genetic alterations'. The former

Abbreviations: 5-FU, 5-fluorouracil; 2′-OMe, 2′-O-methyl; 2′-MOE, 2′-O-methoxyethyl; 3′SS, 3′ splice site; 5′SS, 5′ splice site; BIR, baculovirus IAP repeat; BIRC5, baculoviral IAP repeat-containing 5; CCND1, cyclin D1; CDK, cyclin-dependent kinase; CPT-11, irinotecan hydrochloride; CRC, colorectal cancer; CRNDE, colorectal neoplasia differentially expressed; DPD, dihydropyrimidine dehydrogenase; ESE, exonic splice enhancer; ESS, exonic splice silencer; hnRNP, heterogeneous nuclear ribonucleoprotein; IAP, inhibitor of apoptosis protein; ISE, intronic splice enhancer; ISS, intronic splice silencer; pre-mRNA, precursor messenger RNA; PXR, pregnane X receptor; siRNA, small interfering RNA; SR protein, serine/arginine-rich protein; UGT1A1, UDP glucuronosyltransferase 1A1; VEGF, vascular endothelial growth factor; VEGFR, VEGF receptor.

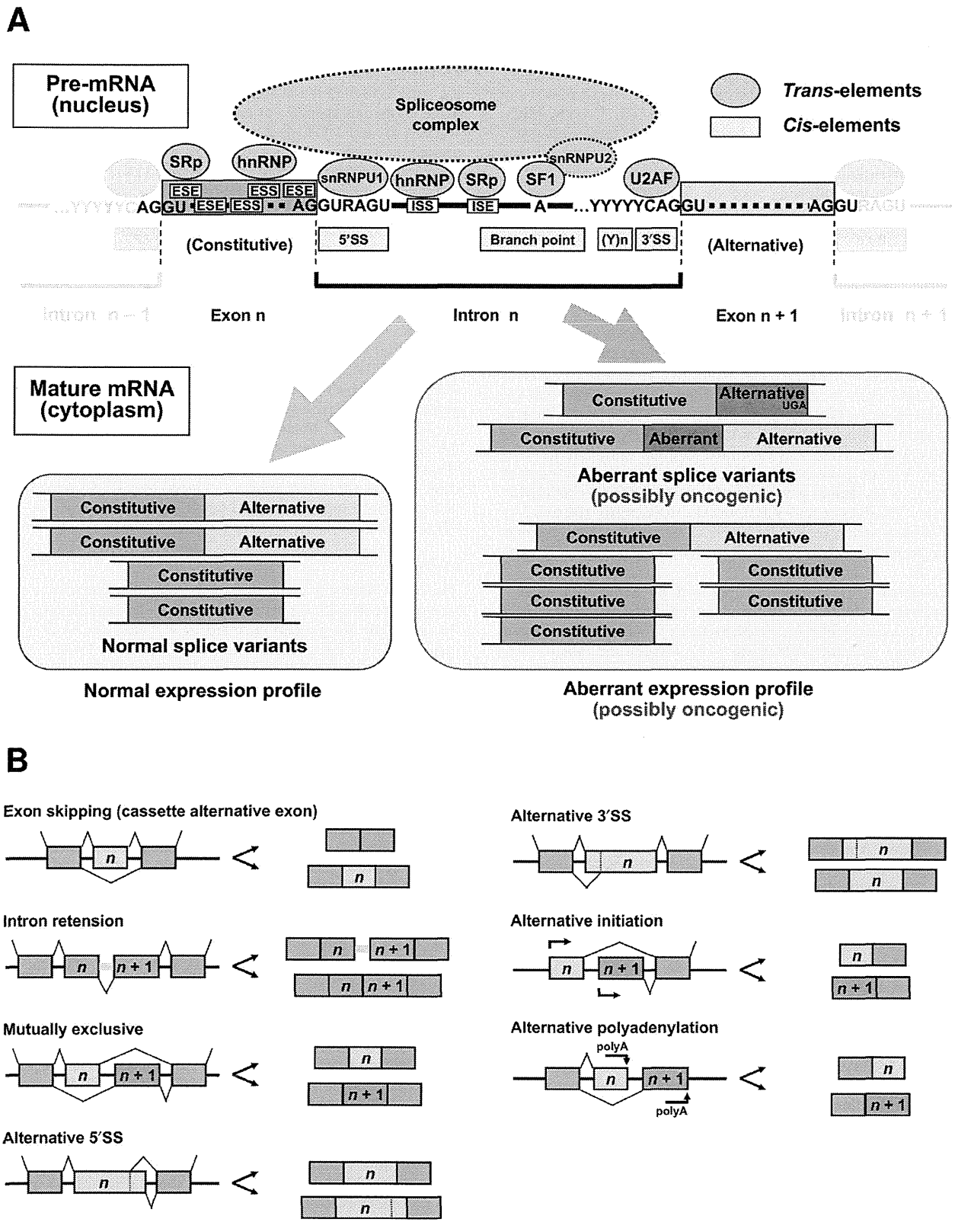

Fig. 1. (**A**) The regulation of alternative pre-mRNA splicing and its alteration in malignancies (adapted from ref. 3). *Cis*-elements and *trans*-elements are indicated with rectangles and ellipses, respectively. In the nucleotide sequences, Y denotes a pyrimidine (U or C) and R denotes a purine (G or A). ESE, exonic splice enhancer; ESS, exonic splice silencer; hnRNP, heterogeneous nuclear ribonucleoprotein; ISE, intronic splice enhancer; ISS, intronic splice silencer; snRNP, small nuclear ribonucleoprotein; SRp, serine/arginine-rich protein; SS, splice site; U2AF, U2 small nuclear ribonucleoprotein auxiliary factor. (**B**) Alternative pre-mRNA splicing (adapted from ref. 3). The green boxes indicate constitutive exons and the blue boxes indicate alternatively spliced exons.

category 'aberrant splice isoforms' can be caused by changes in the 5′-untranslated region (UTR), coding regions and 3′-UTR, as well as ESEs and ESSs, and altered expression of trans-elements may also cause aberrant splice isoforms. In contrast, the latter category 'aberrant expression profiles' can be caused by altered expression and structures of trans-elements, changes in the 5′-UTR, 3′-UTR, ESEs, ESSs, ISEs and ISSs, and possibly by changes of the consensus splice sites in introns. Recent evidence has demonstrated that most splicing occurs cotranscriptionally, and transcription modulates the splicing as well (11). In the following sections, the potential use of splice isoforms as therapeutic targets for CRC, all of which were recently identified, is discussed.

Vascular endothelial growth factor

The vascular endothelial growth factor (*VEGF*) gene superfamily consists of at least six ligands, many of which are spliced to generate a multitude of ligand isoforms (12). The VEGF molecules bind specifically to one or two of the three VEGF receptors (VEGFRs), with VEGFA binding to VEGFR-1 and VEGFR-2 (13). VEGFA and its receptors have been the most common research focus with regard to therapeutic targets with antiangiogenic effects. In addition to the previously known subfamily of VEGFA isoforms (VEGFAxxx), another subfamily, VEGFAxxxb, was identified in 2002 (14). The terminal exon 8 of *VEGFA* gene is spliced in a mutually exclusive manner, resulting in a six amino acid substitution (CDKPRR to SLTRKD) to generate VEGFAxxx and VEGFAxxxb, respectively (Figure 2A). Recent studies have revealed the VEGFAxxx isoform to have proangiogenic properties, whereas VEGFAxxxb has antiangiogenic properties (15,16).

Bevacizumab (17), which was approved for clinical use against CRC in 2004, is a humanized monoclonal antibody that inhibits both the VEGFAxxx and VEGFAxxxb isoforms by blocking their common kinase domain receptor binding site (18). Despite the effectiveness of bevacizumab when it is combined with cytotoxic agents, its low response rate, high rate of resistance and adverse events have been discussed (19). These disadvantages of using bevacizumab may be caused by non-specific targeting, probably resulting from the non-specificity of the antibody for the pro- and antiangiogenic isoforms. In response to these findings, strategies specifically blocking the VEGFAxxx subfamily have been explored and pegaptanib was developed as a short modified RNA aptamer that specifically binds to VEGFAxxx but not VEGFAxxxb (20). Another possible target for CRC is a trans-element, SRp55, which is known to increase VEGFAxxxb expression, leading to antiangiogenic effects (21). In addition, IGF1, TNF-α and TGF-β1 were also identified as being involved in regulating the alternative splicing of VEGFA (21).

UDP glucuronosyltransferase 1A1 and pregnane X receptor

CPT-11, a semisynthetic camptothecin derivative that functions as a topoisomerase I-inhibitor, has been used as an effective anticancer prodrug against CRC. CPT-11 is anabolized to its active metabolite, SN-38, by carboxylesterase (22), and catabolized to its inactive metabolite by glucuronidation (23). UDP glucuronosyltransferase 1A1 (UGT1A1) is the main enzyme involved in glucuronidation of UDP, and genetic polymorphisms of this enzyme, such as UGT1A1*28 (leading to a TA insertion in the promoter region) (24) and single nucleotide polymorphisms (25,26), are known to affect its glucuronidation activity. However, it is widely recognized that the UGT1A1 activity cannot be explained by the polymorphisms alone. One of the main reasons may be the alternative splicing of the *UGT1A* gene. Based on a search of the National Center for Biotechnology Information (NCBI) database and a PubMed literature survey, at least nine isoforms, including UGT1A1, are generated by the alternative splicing of *UGT1A* (Figure 2B), among which UGT1A1, UGT1A7, UGT1A9 and UGT1A10 have glucuronidation activity (27,28), but some of the other isoforms are non-functioning. This means that controlling the alternative splicing of *UGT1A* is important to avoid or decrease the adverse effects associated with CPT-11 and to enhance its efficacy. Recently, Guillemette's group found that the UGT1A locus encodes a previously unknown splice isoform, UGT1A-i2, which is different from the previously known isoform, UGT1A-i1, which results from the alternative splicing of the terminal exon 5 (29). They found that UGT1A-i1 has glucuronosyltransferase activity, but UGT1A-i2 is inactive. By an immunohistochemical analysis using antibodies specific for each of the isoforms, they revealed that UGT1A-i1 and UGT1A-i2 are coproduced in the same structural regions in various organs (30). They further clarified that there is decreased expression of both UGT1A-i1 and UGT1A-i2 in CRC compared with corresponding normal tissues. Knockdown of endogenous UGT1A-i2 enhanced the cellular UGT1A-i1 activity (31), which supports the notion that UGT1A-i2 has a dominant-negative function and is a potential target for regulating the efficacy of CPT-11.

In addition, the UGT1A1 activity is regulated by splice isoforms of the *pregnane X receptor* (*PXR*) gene, which encodes a xenoreceptor that regulates drug metabolism and transporter genes (32). Currently, the *PXR* is known to have three major splice isoforms, T1, T2 and T3 (Figure 2C) (33). The expression of UGT1A1 isoforms, as well as that of UGT1A3 and UGT1A4, is upregulated by the T1 and T2 isoforms, but not by T3, which indicates that splice isoforms of *PXR* are potential therapeutic targets that may regulate the efficacy of CPT-11. In contrast, CPT-11 treatment of HCT116 cells preferentially affected the alternative splicing of factors such as RBM8A, which was not observed in cells treated with cisplatin or vinblastine (34). This indicates that the alternative splicing induced by CPT-11 was not simply due to reduced topoisomerase I activity, but rather was due to rapid RNA polymerase II hyperphosphorylation caused by CPT-11 (34).

Cyclin D1

The cyclin family is composed of proteins that control the progression of the cell cycle by activating cyclin-dependent kinases (CDKs). Among them, the protein encoded by *CCND1* (*cyclin D1*) forms a complex with CDK4 and CDK6. The cyclin D1–CDK4/CDK6 complex induces the phosphorylation of retinoblastoma protein, which releases transcription factors from the phosphorylation of retinoblastoma protein complex, thereby promoting cell division through the G_1-S checkpoint (35). For this reason, *cyclin D1* has been regarded as a proto-oncogene and overexpression of *cyclin D1* occurs at a high frequency in patients with CRC (36,37), esophageal cancer (38) and other malignancies. In addition, *cyclin D1* can activate estrogen receptors in a CDK-independent manner in breast cancer (39) and an abundance of *cyclin D1* affects the radiation sensitivity in some malignancies (40). The transcriptional mechanisms and other functions of cyclin D1 have recently been analyzed (41).

Although genetic alterations of the *cyclin D1* locus are rarely observed, recent studies have demonstrated that the alternative splicing of *cyclin D1* can influence the cancer risk and carcinogenesis (42). The *cyclin D1* gene is known to produce two alternative splice isoforms: CD1a and CD1b (Figure 2D). CD1a is a canonical isoform that consists of five exons, whereas CD1b includes exons 1–4 and a partial intron 4 (43). In colon cancer and other malignancies, the single nucleotide polymorphism G870A, which is the last nucleotide of exon 4 (CC**G** and CC**A**) and is located adjacent to the 5′SS of intron 4 (**GU**RAGU in Figure 1A), modulates the alternative splicing between exon 5 and intron 4, thus generating CD1a and CD1b, respectively (44). In addition, trans-elements ASF/SF2 (45) and Sam68 (46) regulate the alternative splicing toward the generation of CD1b. Although both CD1a and CD1b can associate with CDK4 and CDK6, they show distinct functions and cellular localizations. Phosphorylation of Thr286, which is located within exon 5 (Figure 2D), allows for the nucleocytoplasmic translocalization of *cyclin D1* and its subsequent degradation (47); hence, CD1a can translocate to the cytoplasm, whereas CD1b remains constitutively in the nucleus. Although such functions of CD1a have not been observed, CD1b can cause cellular transformation and has been linked to human carcinogenesis (42,47). By performing the immunocytochemical analyses using antibodies for each of the isoforms, Li *et al.* (40) showed that CD1a, but not CD1b, elicited the DNA damage response in colon cancer cells when stably associated with chromatin. Considering the above results, the

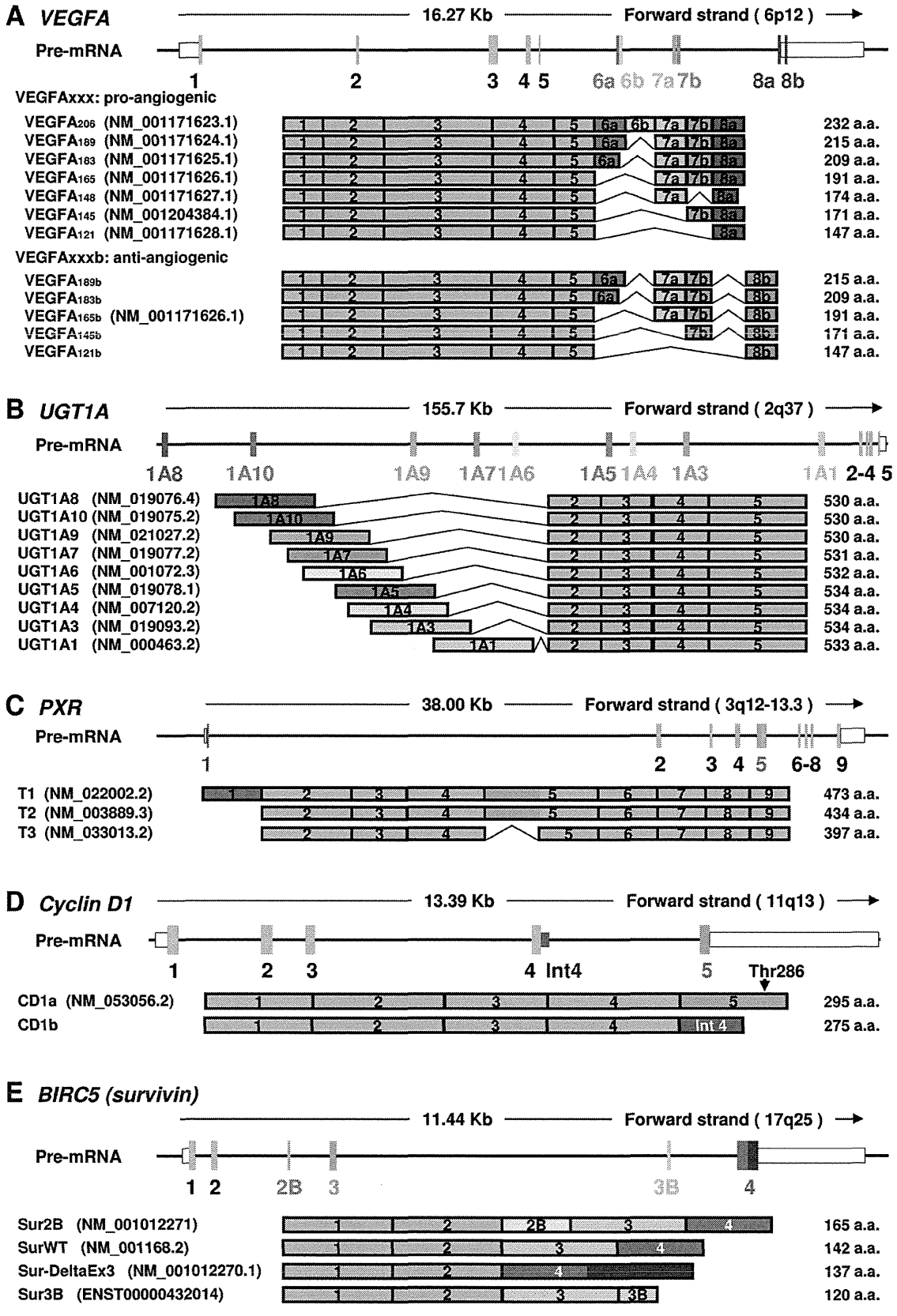

Fig. 2. Splice isoforms of *VEGFA* (**A**), *UGT1A* (**B**), *PXR* (**C**), *cyclin D1* (**D**) and *survivin* (**E**). For each of the genes, the pre-mRNA is indicated at the top and mature mRNAs are indicated below. White boxes indicate 5′-UTR and 3′-UTR. The NM numbers and the numbers of amino acids were provided based on the information contained in the NCBI database (as of 30 September 2012).

two splice isoforms of *cyclin D1* must be distinguished in order to develop therapeutic strategies targeting cyclin D1, and CD1b should be targeted for downregulation to maintain the inherent cell cycle control.

Baculoviral inhibitor of apoptosis protein repeat-containing 5 (survivin)

The inhibitor of apoptosis protein (IAP) family, which is characterized by the presence of baculovirus IAP repeat (BIR) domains (48,49), prevents apoptosis through direct inhibition of caspases and procaspases, and these proteins are expressed at elevated levels in the majority of human malignancies (50). Currently, seven genes in the IAP family have been isolated, among which the baculoviral IAP repeat-containing 5 (BIRC5) gene, also known as survivin, has been the most investigated as a therapeutic target for malignancies, and novel agents targeting this gene or protein are currently under development. Among them, YM155 (Astellas Pharma, Tsukuba, Japan) is a small molecule inhibitor of survivin (51). In 2012, Nakamura *et al.* (51) revealed that YM155 suppresses the expression of survivin through binding to the C-terminal region of interleukin enhancer-binding factor 3, although their study on the molecular mechanism is still underway. LY2181308 (Eli Lilly and Co., Indianapolis, IN) is a second-generation antisense oligonucleotide with a phosphorothioate backbone and other structural modifications, which targets the translation initiation site of the survivin transcripts (52). Both of these agents are designed to block all of the survivin transcripts. Recently, amiloride was reported to regulate the alternative splicing of survivin, as well as that of APAF1 and CRK (53).

Several splice isoforms of survivin have been reported (Figure 2E). In 2007, Sampath and Pelus (54) published a detailed review on the splice isoforms of survivin. The splice isoform Sur2B was regarded to be proapoptotic until the middle of the 2000s. However, the results of recent studies in CRC (55) and other malignancies (56,57) indicated different outcomes. In 2010, Sawai *et al.* (55) reported that Sur2B expression in CRC is an important factor involved in the invasive capacity of tumors in the presence of 5-FU. In 2011, Huang *et al.* (56) reported that the SurWT, Sur-DeltaEx3 and Sur2B isoforms were significantly elevated in astrocytoma and were associated with a poorer prognosis and Vivas-Mejia's study on ovarian cancer cells showed that Sur2B was more abundant in taxane-resistant cells than in taxane-sensitive cells (57). Using CRC samples and corresponding normal tissues, Pavlidou *et al.* (58) analyzed the expression levels of the isoforms, and Antonacopoulou *et al.* (59) analyzed the correlation between the expression of survivin isoforms and single nucleotide polymorphisms. It is still difficult to integrate all of the information on the survivin isoforms because some of the results have been contradictory, but the information will be important to design therapeutic strategies targeting survivin.

Dihydropyrimidine dehydrogenase

After its development in 1957 (60), 5-FU has been a core anticancer agent used for CRC. Approximately 90% of the administered 5-FU is catabolized by dihydropyrimidine dehydrogenase (DPD), mainly in the liver, whereas the remaining 10% of 5-FU is anabolized to exert cytotoxic activity (61), making DPD the most important determinant of 5-FU metabolism (62). Screening for genetic alterations with genomic DNA and mRNA sequencing, van Kuilenburg *et al.* (63) identified genetic alterations in deep intronic regions such as c.1129-5923C>G, which caused aberrant splice isoforms of the *DPD* gene. Their study indicates that caution should be exercised when screening for introns, as well as for exons, when identifying DPD-deficient patients and determining the likely efficacy of 5-FU.

Other splice isoforms implicated in CRC

Some of the other genes encoding splice isoforms that may be possible therapeutic targets for CRC are the *K-RAS* (64), *macroH2A1* (65,66), *SOX9* (67), *SLC39A14* (68,69), colorectal neoplasia differentially expressed (*CRNDE*) (70), *BARD1* (71), *CDH17* (72), *CYP24A1* (73) and *PPARG* genes (74). It is well known that somatic mutation of the *K-RAS* gene is an early event in colorectal carcinogenesis. However, since the middle of the 2000s, the splice isoforms K-RAS4A and K-RAS4B have been reported to have differential functions in apoptosis (75) and differentiation (76) in the intestinal epithelia. In 2009, Abubaker *et al.* (64) analyzed CRC tissues for somatic mutations in the *K-RAS* gene, as well as performing an immunohistochemical analysis of the splice isoforms. Their study demonstrated that the expression of K-RAS4A and K-RAS4B was associated with several clinicopathological features of CRC, and both the *K-RAS* mutation and K-RAS4A expression were independent prognostic markers in a multivariate analysis.

MacroH2A1 is the founding member of the macroH2As family, which has the ability to replace the functions of canonical histones, and has two splice isoforms: macroH2A1.1 and macroH2A1.2. In 2011, Novikov *et al.* (65) demonstrated that the expression of macroH2A1.1 is suppressed in CRC and other malignancies compared with normal tissues. An immunohistochemical study of the two isoforms by Sporn *et al.* (66) in 2012 revealed that the loss of macroH2A1.1 was associated with a worse prognosis of CRC.

The SOX9 transcription factor, which has antioncogenic potential in CRC, generates two isoforms: canonical SOX9 and MiniSOX9, which is a truncated isoform of SOX9 expressed at high levels in CRC (67). An immunohistochemical analysis of CRC and corresponding normal tissues using isoform-specific antibodies revealed that MiniSOX9 behaves as a SOX9 inhibitor and increases the oncogenic potential of CRC cells (67). This indicates that MiniSOX9 may be a therapeutic target for CRC.

SLC39A14 is a divalent cation transporter, which consists of nine exons and has two splice isoforms with a mutually exclusive exon 4, which generates two isoforms: SLC39A14-4A and SLC39A14-4B. In 2011, Thorsen *et al.* (69) demonstrated that SLC39A14-4B mRNA is highly expressed in colonic adenoma and CRC tissues compared with the SLC39A14-4A mRNA. In 2011, Sveen *et al.* (68) reported that SLC39A14-4B can be used as a marker to distinguish CRC from other pathological conditions of the colon. In addition, Graham *et al.* (70) indicated that splice isoforms of the *CRNDE* gene seem to be differentially expressed in different stages of CRC.

In 2011, Yi and Tang reported a review article on the potential use of splice variants as diagnostic, predictive and prognostic markers for CRC (77), which included information about *APC*, *TIMP-1*, *VEGFA*, *DYX1C1* and *c-FLIP*, among other genes. Their article provided information about the splice isoforms with regard to the use of chemotherapy for CRC. A genome-wide exon array analysis in 2011 detected several CRC-specific splice isoforms (*TCF12*, *OSBPL1A*, *TRAK1*, *ANK3*, *CHEK1*, *UGP2*, *LMO7*, *ACSL5* and *SCIN*) (78). In addition, trans-elements, such as SR protein kinase 1 and SR protein kinase 2, have also been discussed as therapeutic targets for CRC and other malignancies (79), although trans-elements are not described in detail in this review. Most of the studies presented here were reported after 2011, and the information on splice isoforms is still being accumulated.

Therapeutic strategies to target splice isoforms

Therapeutic targeting of splice isoforms may be achieved through conventional small molecules, but these molecules can only target a small subset of proteins, such as enzymes (e.g. tyrosine kinases) and receptors (e.g. the epidermal growth factor receptor). On the other hand, RNA-based therapeutics can theoretically target all of the pre-mRNAs and mRNAs with a wider range and higher selectivity than small molecules (80), although almost all of these modalities are still in preclinical development. Currently, the most important issue to be resolved for the use of RNA-based therapeutics as macromolecules is the development of an optimal drug delivery system.

The RNA-based therapeutics include antisense oligonucleotides, small interfering RNA (siRNA), splice-switching oligonucleotides and other molecules such as ribozymes and aptamers. Among them, synthetically modified antisense oligonucleotides are about 20 nucleotides long (Figure 3A) and the annealing of the oligonucleotides

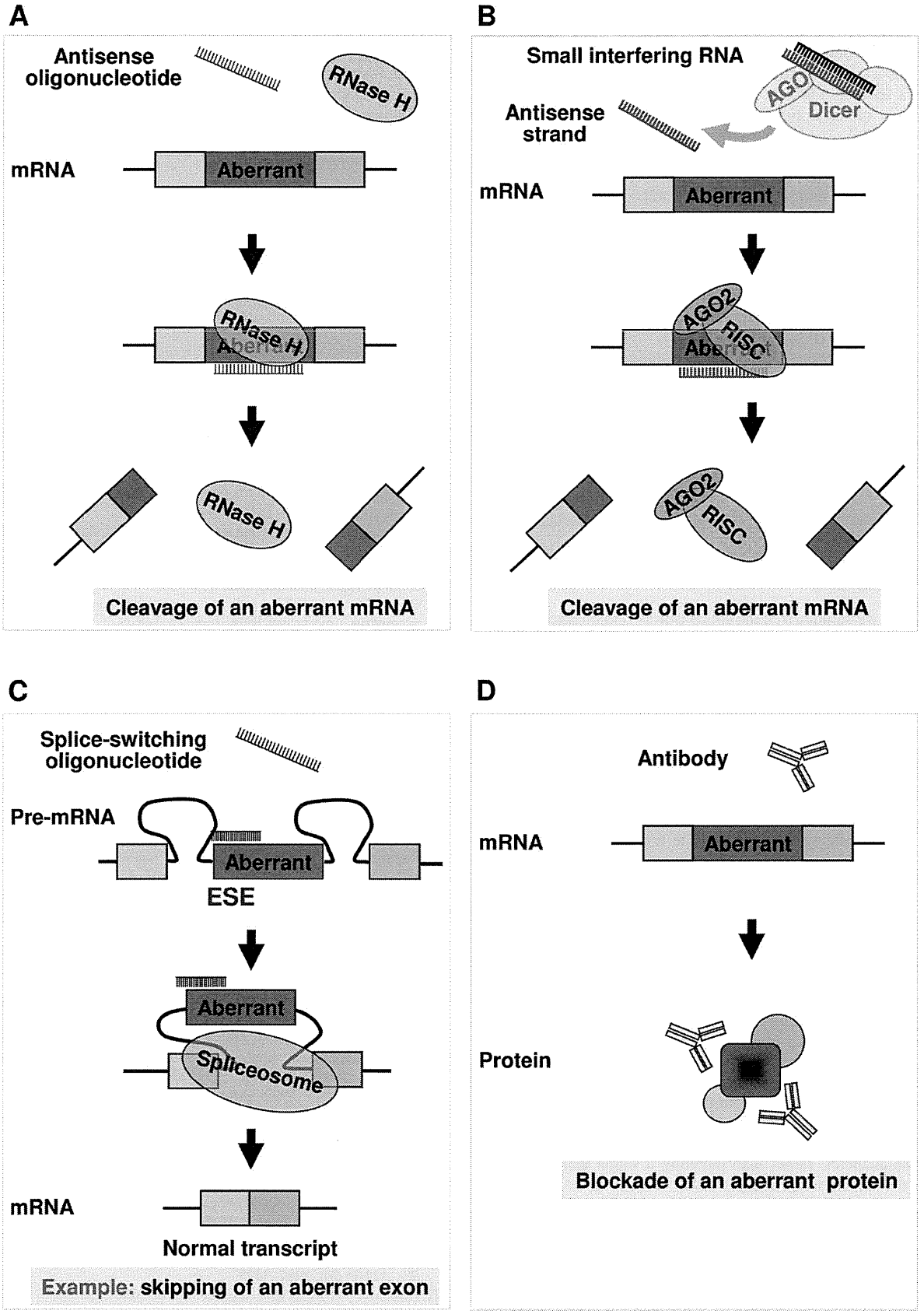

Fig. 3. Macromolecules targeting splice isoforms associated with malignancy. (**A**) A synthetically modified antisense oligonucleotide, (**B**) siRNA, (**C**) a splice-switching oligonucleotide and (**D**) an antibody. In (**C**), an example of splice switching is presented, in which an ESE located in the aberrant exon is annealed to an oligonucleotide, and the aberrant exon is skipped. AGO, argonaute; ESE, exonic splice enhancer; RISC, RNA-induced silencing complex; RNase H, ribonuclease H.

to mRNA allows the cleavage of the mRNA by ribonuclease H. To provide enhanced structural stability and pharmacological qualities while not interfering with the activity of ribonuclease H, various modifications of their chemical structures have been made, such as the use of a phosphorothioate linkage instead of natural phosphates as a backbone of nucleotides, and 2′-O-methyl (2′-OMe) residues, 2′-O-methoxyethyl (2′-MOE) residues or locked nucleic acids (81) have also been developed. In addition, the chemistry-dependent toxicities induced by their structures is another important issue to be considered. siRNA (Figure 3B) is another modality that consists of a double-stranded RNA fragment 21–22 nucleotides long. After interacting with the multiprotein RNA-induced silencing complex, the antisense strand of the siRNA anneals to the complementary mRNA as a target, and the endonuclease argonaute 2 cleaves the annealed mRNA. In this modality, off-target effects and the innate immune response via the activation of Toll-like receptors should be carefully managed. Splice-switching oligonucleotides (Figure 3C) modulate pre-mRNA splicing with spliceosomes and repair the defective pre-mRNA to generate proteins that have distinct functions. Monoclonal antibodies (82) have also been used for various targets (Figure 3D). In contrast with the RNA-based therapeutics, some antibodies targeting oncogenic proteins have already been in clinical use; and those for each of the splice isoforms will likely be further developed for clinical use. The development of antibodies is still very expensive, and further considerations for their development are discussed elsewhere (83). Targeting trans-elements that act as spliceosomes or splicing modulators is another option.

Future perspectives

In this review, we summarized the splice isoforms that represent possible therapeutic targets for CRC. As discussed in this review, isoform-specific antibodies for VEGFA (16), UGT1A (30), cyclin D1 (40), K-RAS (64) and SOX9 (67) are currently available, and they can be utilized for immunohistochemical analyses and other purposes, and may eventually be useful for clinical applications. With the recent advances in nucleotide sequencing technologies, an entire set of genomic DNA sequences has been analyzed, and in the next stage, an entire set of RNA sequences will be further analyzed; the interpretation of the latter, however, is far more complex compared with the former, mainly due to the wide variety of mature mRNAs resulting from alternative splicing. To elucidate the regulatory mechanism(s) for alternative splicing as a whole, the two sets of sequence information will have to be integrated. Although the importance of individual cis-elements in the splicing machinery has been widely discussed, the concept of a 'splicing code', which is defined as a complex combination of the cis-elements that direct constitutive or alternative splicing, was proposed as early as the 1970s. To experimentally prove this concept had been a major challenge, but recent studies combining transcriptome-wide data with advanced machine learning algorithms were able to predict new classes of alternative splicing events under regulation by the splicing code (84,85). Furthermore, in a recent genome-wide siRNA screening, Moore *et al.* (6) identified a coordinated alternative splicing of *Bcl-X*, *MCL1*, *CASP9* and other apoptosis-associated genes under a common regulatory network. These findings suggest that we should consider a set of splice isoforms or their common regulatory network when developing therapeutic strategies for malignancies, rather than targeting a single gene. To what extent the mechanisms regulating alternative splicing are organ-specific remains unclear, but the phenomenon is complex, and is the subject of many ongoing studies.

Research in these various areas is still ongoing, and new discoveries are being reported at a rapid pace. Recent reports have demonstrated that alternative splicing is also affected by newly identified regulatory factors, such as RNA polymerase II elongation (86), the chromatin structure (87,88), histone modifications (89), the RNA structure (90) and the spliceosome structure (91), most of which are interwoven bidirectionally (87,88). Importantly, the splicing machinery is regulated by innate microRNAs, siRNAs, small nucleolar RNAs and other non-coding RNAs (92,93), and these should also be considered as therapeutic targets. The rapidly increasing information available about nucleotide sequences, trans-elements, and newly identified regulatory factors, along with novel bioinformatics technology, such as the multi-mapping Bayesian gene eXpression (MMBGX) program by Turro *et al.* (94), which enables the detection of differential splicing at the isoform level, will provide additional information about how therapeutic strategies targeting alternative splicing in malignancies can be developed.

Funding

HIROMI Medical Research Foundation (Sendai, Japan); Osaka Basic Medical Research Promoting Foundation (Osaka, Japan).

Conflict of Interest Statement: None declared.

References

1. Jemal,A. *et al.* (2011) Global cancer statistics. *CA. Cancer J. Clin.*, **61**, 69–90.
2. http://www.nccn.org/professionals/physician_gls/f_guidelines.asp#site
3. Miura,K. *et al.* (2011) Alternative pre-mRNA splicing in digestive tract malignancy. *Cancer Sci.*, **102**, 309–316.
4. Wang,E.T. *et al.* (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470–476.
5. Yoshida,K. *et al.* (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, **478**, 64–69.
6. Moore,M.J. *et al.* (2010) An alternative splicing network links cell-cycle control to apoptosis. *Cell*, **142**, 625–636.
7. Chow,L.T. *et al.* (1977) An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, **12**, 1–8.
8. Klessig,D.F. (1977) Two adenovirus mRNAs have a common 5' terminal leader sequence encoded at least 10kb upstream from their main coding regions. *Cell*, **12**, 9–21.
9. Chen,M. *et al.* (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.*, **10**, 741–754.
10. Warzecha,C.C. *et al.* (2009) ESRP1 and ESRP2 are epithelial cell-type-specific regulators of FGFR2 splicing. *Mol. Cell*, **33**, 591–601.
11. Luco,R.F. *et al.* (2011) Epigenetics in alternative pre-mRNA splicing. *Cell*, **144**, 16–26.
12. Amankwah,E.K. *et al.* (2012) Gene variants in the angiogenesis pathway and prostate cancer. *Carcinogenesis*, **33**, 1259–1269.
13. Rennel,E.S. *et al.* (2009) Therapeutic potential of manipulating VEGF splice isoforms in oncology. *Future Oncol.*, **5**, 703–712.
14. Bates,D.O. *et al.* (2002) VEGF165b, an inhibitory splice variant of vascular endothelial growth factor, is down-regulated in renal cell carcinoma. *Cancer Res.*, **62**, 4123–4131.
15. Harper,S.J. *et al.* (2008) VEGF-A splicing: the key to anti-angiogenic therapeutics? *Nat. Rev. Cancer*, **8**, 880–887.
16. Rennel,E.S. *et al.* (2009) VEGF(121)b, a new member of the VEGF(xxx)b family of VEGF-A splice isoforms, inhibits neovascularisation and tumour growth *in vivo*. *Br. J. Cancer*, **101**, 1183–1193.
17. Presta,L.G. *et al.* (1997) Humanization of an anti-vascular endothelial growth factor monoclonal antibody for the therapy of solid tumors and other disorders. *Cancer Res.*, **57**, 4593–4599.
18. Muller,Y.A. *et al.* (1997) Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 7192–7197.
19. Troiani,T. *et al.* (2012) Beyond bevacizumab: new anti-VEGF strategies in colorectal cancer. *Expert Opin. Investig. Drugs*, **21**, 949–959.
20. Gragoudas,E.S. *et al.*; VEGF Inhibition Study in Ocular Neovascularization Clinical Trial Group. (2004) Pegaptanib for neovascular age-related macular degeneration. *N. Engl. J. Med.*, **351**, 2805–2816.
21. Nowak,D.G. *et al.* (2008) Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors. *J. Cell. Sci.*, **121**(Pt 20), 3487–3495.
22. Haaz,M.C. *et al.* (1998) Metabolism of irinotecan (CPT-11) by human hepatic microsomes: participation of cytochrome P-450 3A and drug interactions. *Cancer Res.*, **58**, 468–472.
23. Rivory,L.P. *et al.* (1997) Pharmacokinetic interrelationships of irinotecan (CPT-11) and its three major plasma metabolites in patients enrolled in phase I/II trials. *Clin. Cancer Res.*, **3**, 1261–1266.

24. Ando,Y. *et al.* (2000) Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis. *Cancer Res.*, **60**, 6921–6926.
25. Sai,K. *et al.* (2004) UGT1A1 haplotypes associated with reduced glucuronidation and increased serum bilirubin in irinotecan-administered Japanese patients with cancer. *Clin. Pharmacol. Ther.*, **75**, 501–515.
26. Innocenti,F. *et al.* (2004) Genetic variants in the UDP-glucuronosyltransferase 1A1 gene predict the risk of severe neutropenia of irinotecan. *J. Clin. Oncol.*, **22**, 1382–1388.
27. Gong,Q.H. *et al.* (2001) Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics*, **11**, 357–368.
28. Ritter,J.K. *et al.* (1992) A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *J. Biol. Chem.*, **267**, 3257–3261.
29. Girard,H. *et al.* (2007) Genetic diversity at the UGT1 locus is amplified by a novel 3' alternative splicing mechanism leading to nine additional UGT1A proteins that act as regulators of glucuronidation activity. *Pharmacogenet. Genomics*, **17**, 1077–1089.
30. Bellemare,J. *et al.* (2011) Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans. *J. Pathol.*, **223**, 425–435.
31. Bellemare,J. *et al.* (2010) Alternatively spliced products of the UGT1A gene interact with the enzymatically active proteins to inhibit glucuronosyltransferase activity *in vitro*. *Drug Metab. Dispos.*, **38**, 1785–1789.
32. Raynal,C. *et al.* (2010) Pregnane X Receptor (PXR) expression in colorectal cancer cells restricts irinotecan chemosensitivity through enhanced SN-38 glucuronidation. *Mol. Cancer*, **9**, 46.
33. Gardner-Stephen,D. *et al.* (2004) Human PXR variants and their differential effects on the regulation of human UDP-glucuronosyltransferase gene expression. *Drug Metab. Dispos.*, **32**, 340–347.
34. Solier,S. *et al.* (2010) Genome-wide analysis of novel splice variants induced by topoisomerase I poisoning shows preferential occurrence in genes encoding splicing factors. *Cancer Res.*, **70**, 8055–8065.
35. Lew,D.J. *et al.* (1991) Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. *Cell*, **66**, 1197–1206.
36. Arber,N. *et al.* (1996) Increased expression of cyclin D1 is an early event in multistage colorectal carcinogenesis. *Gastroenterology*, **110**, 669–674.
37. Wang,Q.S. *et al.* (1998) Altered expression of cyclin D1 and cyclin-dependent kinase 4 in azoxymethane-induced mouse colon tumorigenesis. *Carcinogenesis*, **19**, 2001–2006.
38. Jiang,W. *et al.* (1993) Altered expression of the cyclin D1 and retinoblastoma genes in human esophageal cancer. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 9026–9030.
39. Zwijsen,R.M. *et al.* (1997) CDK-independent activation of estrogen receptor by cyclin D1. *Cell*, **88**, 405–415.
40. Li,Z. *et al.* (2010) Alternative cyclin D1 splice forms differentially regulate the DNA damage response. *Cancer Res.*, **70**, 8802–8811.
41. Witzel,I.I. *et al.* (2010) Regulation of cyclin D1 gene expression. *Biochem. Soc. Trans.*, **38**, 217–222.
42. Knudsen,K.E. (2006) The cyclin D1b splice variant: an old oncogene learns new tricks. *Cell Div.*, **1**, 15.
43. Betticher,D.C. *et al.* (1995) Alternate splicing produces a novel cyclin D1 transcript. *Oncogene*, **11**, 1005–1011.
44. Knudsen,K.E. *et al.* (2006) Cyclin D1: polymorphism, aberrant splicing and cancer risk. *Oncogene*, **25**, 1620–1628.
45. Olshavsky,N.A. *et al.* (2010) Identification of ASF/SF2 as a critical, allele-specific effector of the cyclin D1b oncogene. *Cancer Res.*, **70**, 3975–3984.
46. Paronetto,M.P. *et al.* (2010) Alternative splicing of the cyclin D1 proto-oncogene is regulated by the RNA-binding protein Sam68. *Cancer Res.*, **70**, 229–239.
47. Solomon,D.A. *et al.* (2003) Cyclin D1 splice variants. Differential effects on localization, RB phosphorylation, and cellular transformation. *J. Biol. Chem.*, **278**, 30339–30347.
48. Hinds,M.G. *et al.* (1999) Solution structure of a baculoviral inhibitor of apoptosis (IAP) repeat. *Nat. Struct. Biol.*, **6**, 648–651.
49. Sun,C. *et al.* (1999) NMR structure and mutagenesis of the inhibitor-of-apoptosis protein XIAP. *Nature*, **401**, 818–822.
50. Miura,K. *et al.* (2009) cIAP2 as a therapeutic target in colorectal cancer and other malignancies. *Expert Opin. Ther. Targets*, **13**, 1333–1345.
51. Nakamura,N. *et al.* (2012) ILF3/NF110 is a target of YM155, a suppressant of survivin. *Mol. Cell. Proteomics*, **11**, M111.013243.
52. Carrasco,R.A. *et al.* (2011) Antisense inhibition of survivin expression as a cancer therapeutic. *Mol. Cancer Ther.*, **10**, 221–232.
53. Chang,W.H. *et al.* (2011) Amiloride modulates alternative splicing in leukemic cells and resensitizes Bcr-AblT315I mutant cells to imatinib. *Cancer Res.*, **71**, 383–392.
54. Sampath,J. *et al.* (2007) Alternative splice variants of survivin as potential targets in cancer. *Curr. Drug Discov. Technol.*, **4**, 174–191.
55. Sawai,K. *et al.* (2010) Survivin-3B gene decreases the invasion-inhibitory effect of colon cancer cells with 5-fluorouracil. *Oncol. Res.*, **18**, 541–547.
56. Huang,Y. *et al.* (2011) Expression and prognostic significance of survivin splice variants in diffusely infiltrating astrocytoma. *J. Clin. Pathol.*, **64**, 953–959.
57. Vivas-Mejia,P.E. *et al.* (2011) Silencing survivin splice variant 2B leads to antitumor activity in taxane–resistant ovarian cancer. *Clin. Cancer Res.*, **17**, 3716–3726.
58. Pavlidou,A. *et al.* (2011) Survivin isoforms and clinicopathological characteristics in colorectal adenocarcinomas using real-time qPCR. *World J. Gastroenterol.*, **17**, 1614–1621.
59. Antonacopoulou,A.G. *et al.* (2011) The survivin -31 snp in human colorectal cancer correlates with survivin splice variant expression and improved overall survival. *Cell. Oncol. (Dordr).*, **34**, 381–391.
60. Heidelberger,C. *et al.* (1957) Fluorinated pyrimidines, a new class of tumour-inhibitory compounds. *Nature*, **179**, 663–666.
61. Heggie,G.D. *et al.* (1987) Clinical pharmacokinetics of 5-fluorouracil and its metabolites in plasma, urine, and bile. *Cancer Res.*, **47**, 2203–2206.
62. Miura,K. *et al.* (2012) S-1 as a core anticancer fluoropyrimidine agent. *Expert Opin. Drug Deliv.*, **9**, 273–286.
63. van Kuilenburg,A.B. *et al.* (2010) Intragenic deletions and a deep intronic mutation affecting pre-mRNA splicing in the dihydropyrimidine dehydrogenase gene as novel mechanisms causing 5-fluorouracil toxicity. *Hum. Genet.*, **128**, 529–538.
64. Abubaker,J. *et al.* (2009) Prognostic significance of alterations in KRAS isoforms KRAS-4A/4B and KRAS mutations in colorectal carcinoma. *J. Pathol.*, **219**, 435–445.
65. Novikov,L. *et al.* (2011) QKI-mediated alternative splicing of the histone variant MacroH2A1 regulates cancer cell proliferation. *Mol. Cell Biol.*, **31**, 4244–4255.
66. Sporn,J.C. *et al.* (2012) Differential regulation and predictive potential of MacroH2A1 isoforms in colon cancer. *Am. J. Pathol.*, **180**, 2516–2526.
67. Abdel-Samad,R. *et al.* (2011) MiniSOX9, a dominant-negative variant in colon cancer cells. *Oncogene*, **30**, 2493–2503.
68. Sveen,A. *et al.* (2012) The exon-level biomarker SLC39A14 has organ-confined cancer-specificity in colorectal cancer. *Int. J. Cancer*, **131**, 1479–1485.
69. Thorsen,K. *et al.* (2011) Alternative splicing of SLC39A14 in colorectal cancer is regulated by the Wnt pathway. *Mol. Cell Proteomics*, **10**, M110.002998.
70. Graham,L.D. *et al.* (2011) Colorectal neoplasia differentially expressed (CRNDE), a novel gene with elevated expression in colorectal adenomas and adenocarcinomas. *Genes Cancer*, **2**, 829–840.
71. Sporn,J.C. *et al.* (2011) BARD1 expression predicts outcome in colon cancer. *Clin. Cancer Res.*, **17**, 5451–5462.
72. Lee,N.P. *et al.* (2010) Role of cadherin-17 in oncogenesis and potential therapeutic implications in hepatocellular carcinoma. *Biochim. Biophys. Acta*, **1806**, 138–145.
73. Horváth,H.C. *et al.* (2010) CYP24A1 splice variants–implications for the antitumorigenic actions of 1,25-(OH)2D3 in colorectal cancer. *J. Steroid Biochem. Mol. Biol.*, **121**, 76–79.
74. Koh,W.P. *et al.* (2006) Peroxisome proliferator-activated receptor (PPAR) gamma gene polymorphisms and colorectal cancer risk among Chinese in Singapore. *Carcinogenesis*, **27**, 1797–1802.
75. Plowman,S.J. *et al.* (2006) The K-Ras 4A isoform promotes apoptosis but does not affect either lifespan or spontaneous tumor incidence in aging mice. *Exp. Cell Res.*, **312**, 16–26.
76. Patek,C.E. *et al.* (2008) The pro-apoptotic K-Ras 4A proto-oncoprotein does not affect tumorigenesis in the ApcMin/+ mouse small intestine. *BMC Gastroenterol.*, **8**, 24.
77. Yi,Q. *et al.* (2011) Alternative spliced variants as biomarkers of colorectal cancer. *Curr. Drug Metab.*, **12**, 966–974.
78. Thorsen,K. *et al.* (2011) Tumor-specific usage of alternative transcription start sites in colorectal cancer identified by genome-wide exon array analysis. *BMC Genomics*, **12**, 505.
79. Hayes,G.M. *et al.* (2007) Serine-arginine protein kinase 1 overexpression is associated with tumorigenic imbalance in mitogen-activated protein kinase pathways in breast, colonic, and pancreatic carcinomas. *Cancer Res.*, **67**, 2072–2080.
80. Kole,R. *et al.* (2012) RNA therapeutics: beyond RNA interference and antisense oligonucleotides. *Nat. Rev. Drug Discov.*, **11**, 125–140.
81. Deleavey,G.F. *et al.* (2012) Designing chemically modified oligonucleotides for targeted gene silencing. *Chem. Biol.*, **19**, 937–954.
82. Pillay,V. *et al.* (2011) Antibodies in oncology. *N. Biotechnol.*, **28**, 518–529.

83. Tabrizi,M.A. *et al.* (2009) Translational strategies for development of monoclonal antibodies from discovery to the clinic. *Drug Discov. Today*, **14**, 298–305.
84. Barash,Y. *et al.* (2010) Deciphering the splicing code. *Nature*, **465**, 53–59.
85. Irimia,M. *et al.* (2012) Alternative splicing: decoding an expansive regulatory layer. *Curr. Opin. Cell Biol.*, **24**, 323–332.
86. Montes,M. *et al.* (2012) Functional coupling of transcription and splicing. *Gene*, **501**, 104–117.
87. Luco,R.F. *et al.* (2011) More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. *Curr. Opin. Genet. Dev.*, **21**, 366–372.
88. de Almeida,S.F. *et al.* (2012) Design principles of interconnections between chromatin and pre-mRNA splicing. *Trends Biochem. Sci.*, **37**, 248–253.
89. Luco,R.F. *et al.* (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
90. McManus,C.J. *et al.* (2011) RNA structure and the mechanisms of alternative splicing. *Curr. Opin. Genet. Dev.*, **21**, 373–379.
91. Will,C.L. *et al.* (2011) Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.*, **3**, pii: a003707.
92. Gromak,N. (2012) Intronic microRNAs: a crossroad in gene regulation. *Biochem. Soc. Trans.*, **40**, 759–761.
93. Esteller,M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.*, **12**, 861–874.
94. Turro,E. *et al.* (2010) MMBGX: a method for estimating expression at the isoform level and detecting differential splicing using whole-transcript Affymetrix arrays. *Nucleic Acids Res.*, **38**, e4.

Received August 2, 2012; revised October 10, 2012; accepted October 25, 2012