Cancer Letters xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Mini-review

Fusion genes and their discovery using high throughput sequencing

M.J. Annala [a,*], B.C. Parker [b], W. Zhang [b], M. Nykter [a]

[a] Tampere University of Technology, Tampere, Finland
[b] The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Fusion gene
High throughput sequencing
Cancer

ABSTRACT

Fusion genes are hybrid genes that combine parts of two or more original genes. They can form as a result of chromosomal rearrangements or abnormal transcription, and have been shown to act as drivers of malignant transformation and progression in many human cancers. The biological significance of fusion genes together with their specificity to cancer cells has made them into excellent targets for molecular therapy. Fusion genes are also used as diagnostic and prognostic markers to confirm cancer diagnosis and monitor response to molecular therapies. High-throughput sequencing has enabled the systematic discovery of fusion genes in a wide variety of cancer types. In this review, we describe the history of fusion genes in cancer and the ways in which fusion genes form and affect cellular function. We also describe computational methodologies for detecting fusion genes from high-throughput sequencing experiments, and the most common sources of error that lead to false discovery of fusion genes.

1. Introduction

1.1. Fusion genes in cancer

Somatic fusion genes are regarded as one of the major drivers behind cancer initiation and progression (reviewed in [1]). The first signs of fusion genes in human cancer were identified in 1960 when a reciprocal translocation between the q-arms of chromosomes 9 and 22 was discovered in over 95% of chronic myelogenous leukemia patients [2,3]. After two decades the translocation was understood to produce a chimeric *BCR-ABL1* transcript that encoded a constitutively active form of the ABL kinase [4]. At the same time, Burkitt's lymphoma was found to harbor activating fusions between immunoglobulin genes and *MYC* [5–7]. These initial findings were promptly followed by the discovery of dozens of new fusion genes in human cancers (Table 1). Among hematological malignancies, the identification of a *PML-RARA* fusion in acute promyelocytic leukemia paved the way for an effective tretinoin-based molecular therapy [8,9], while a *RUNX1-ETO* chimeric protein was found to characterize a subtype of acute myeloid leukemia with prolonged median survival [10]. Success stories among solid cancers included the early discovery of fusions between *EWSR1* and members of the *ETS* transcription factor family in Ewing's sarcoma [11,12], and the discovery of characteristic *SS18-SSX* fusions in synovial sarcoma [13–15]. In myxoid liposarcoma, *FUS-DDIT3* and *EWSR1-DDIT3* fusions were found to be pathognomonic for the disease [16–18]. Despite these discoveries, fusion positive cases only accounted for a tiny fraction of all solid cancers. This changed in 2005 when fusion genes juxtaposing *TMPRSS2* and members of the *ETS* transcription factor family were found in 70% of prostate cancers [19]. Subsequent discoveries in solid cancers included the discovery of *EML4-ALK* fusions and *CHD7* rearrangements in non-small cell lung cancer [20–22], *KIAA1549-BRAF* fusions in pediatric glioma [23], *FGFR3–TACC3* fusions in glioblastoma [24,25], and R-spondin fusions in colon cancer [26]. Some cancers were found to associate with multiple fusion genes that presented in a mutually exclusive manner. For instance, the fusions *TMPRSS2-ERG* and *TMPRSS2-ETV1* are common in prostate cancer, but almost never co-occur in a single tumor [19]. Similarly, the fusion genes SS18-SSX1 and SS18-SSX2 are found in 70% and 30% of synovial sarcoma patients, but never co-occur [27]. In some cases, fusion genes also exhibit mutual exclusivity or co-occurrence with other types of genomic aberrations, as exemplified by the mutual exclusivity of *ETS* fusions and *SPINK1* overexpression in prostate cancer [28]. Mutual exclusivity between two genomic alterations usually implies that the two alterations confer similar contributions to the malignant phenotype, and therefore oncogenic selection ceases after one alteration has been acquired.

Some fusion genes are found recurrently in multiple cancers. The *BCR-ABL1* fusion gene is recurrent in both chronic myelogenous leukemia [3] and acute lymphocytic leukemia [29], and isolated cases have been reported in other leukemias. *TPM3-ALK* fusions provide an example of a fusion gene found in cancer cells of completely different lineages. *TPM3-ALK* is found in 15% of cases of anaplastic large cell lymphoma, a hematological malignancy of T-cell origin [30], and in 50% of inflammatory myofibroblastic tu-

* Corresponding author.
E-mail address: matti.annala@tut.fi (M.J. Annala).

0304-3835/$ - see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.canlet.2013.01.011

Table 1
Fusion genes in human cancers.

	Cancer	Fusion gene	Frequency (%)	Mechanism of formation	Biological impact	References
Hematological cancers	Acute lymphocytic leukemia	ETV6-RUNX1	25	Interchromosomal translocation	Oncogenic chimeric protein	Golub et al. (1995) and Romana et al. (1995)
		BCR-ABL1	15	Interchromosomal translocation	Oncogenic chimeric protein	Westbrook et al. (1992)
	Acute myeloid leukemia	RUNX1-ETO	10–15	Interchromosomal translocation	Oncogenic chimeric protein	Erickson et al. (1992)
		CBFB-MYH11	10–15	Inversion	Oncogenic chimeric protein	Liu et al. (1993)
	Acute promyelocytic leukemia	PML-RARA	95	Interchromosomal translocation	Oncogenic chimeric protein	Borrow et al. (1990) and Warrell et al. (1991)
		PLZF-RARA	0–5	Interchromosomal translocation	Oncogenic chimeric protein	Chen et al. (1993)
	Anaplastic large cell lymphoma	NPM1-ALK	75	Interchromosomal translocation	Oncogenic chimeric protein	Morris et al. (1994)
		TPM3-ALK	15	Interchromosomal translocation	Oncogenic chimeric protein	Lamant et al. (1999)
	Burkitt's lymphoma	IG@-MYC	90–100	Interchromosomal translocation	Promoter exchange	Manolov et al. (1972) and Dalla-Favera et al. (1982)
	Chronic myelogenous leukemia	BCR-ABL1	95–100	Interchromosomal translocation	Oncogenic chimeric protein	Nowell et al. (1960) and Shtivelman et al. (1985)
Solid cancers	Adenoid cystic carcinoma	MYB-NFIB	90–100	Interchromosomal translocation	Loss of microRNA regulation	Persson et al. (2009)
	Bladder cancer	FGFR3–TACC3	0–10	Tandem duplication	Oncogenic chimeric protein	Williams et al. (2012)
	Clear cell sarcoma	EWSR1-ATF1	90–100	Interchromosomal translocation	Oncogenic chimeric protein	Bridge et al. (1990) and Zucman et al. (1993)
	Colon cancer	PTPRK-RSPO3	5–10	Inversion	Promoter exchange	Seshagiri et al. (2012)
		EIF3E3-RSPO2	0–5	Deletion	Promoter exchange	Seshagiri et al. (2012)
	Congenital fibrosarcoma	ETV6-NTRK3	90–100	Interchromosomal translocation	Oncogenic chimeric protein	Knezevich et al. (1998)
	Ewing sarcoma	EWSR1-FLI1	90	Interchromosomal translocation	Oncogenic chimeric protein	Turc-Carel et al. (1983) and Aurias et al. (1983)
	Follicular thyroid carcinoma	PAX8-PPARG	60	Interchromosomal translocation	Oncogenic chimeric protein	Kroll et al. (2000)
	Glioblastoma	FGFR3–TACC3	0–5	Tandem duplication	Oncogenic chimeric protein	Singh et al. (2012) and Parker et al. (2012)
	Inflammatory myofibroblastic tumor	TPM3-ALK	50	Interchromosomal translocation	Oncogenic chimeric protein	Lawrence et al. (2000)
	Mucoepidermoid carcinoma	MECT1-MAML2	60	Interchromosomal translocation	Oncogenic chimeric protein	Tonon et al. (2003)
	Myxoid liposarcoma	FUS-DDIT3	90–100	Interchromosomal translocation	Oncogenic chimeric protein	Crozat et al. (1993) and Rabbits et al. (1993)
		EWSR1-DDIT3	0–5	Interchromosomal translocation	Oncogenic chimeric protein	Panagopoulos et al. (1996)
	Non-small cell lung cancer	EML4-ALK	0–10	Inversion	Oncogenic chimeric protein	Soda et al. (2007) and Rikova et al. (2007)
	NUT midline carcinoma	BRD4-NUT	90–100	Interchromosomal translocation	Promoter exchange	French et al. (2003)
	Papillary thyroid carcinoma	CCDC6-RET	15	Inversion	Oncogenic chimeric protein	Grieco et al. (1990)
		NCOA4-RET	15	Complex rearrangement	Oncogenic chimeric protein	Santoro et al. (1994)
	Pediatric renal cell carcinoma	PRCC-TFE3	20–40	Interchromosomal translocation	Oncogenic chimeric protein	Weterman et al. (1996)
	Pilocytic astrocytoma	KIAA1549-BRAF	70	Tandem duplication	Oncogenic chimeric protein	Jones et al. (2008)
	Prostate cancer	TMPRSS2-ERG	60	Deletion	Promoter exchange	Tomlins et al. (2005)
		TMPRSS2-ETV1	0–5	Interchromosomal translocation	Promoter exchange	Tomlins et al. (2005)
		TMPRSS2-ETV4	0–5	Interchromosomal translocation	Promoter exchange	Tomlins et al. (2006)
	Secretory breast carcinoma	ETV6-NTRK3	90	Interchromosomal translocation	Oncogenic chimeric protein	Tognon et al. (2002)
	Serous ovarian cancer	ESRRA-C11orf20	15	Intrachromosomal translocation	Oncogenic chimeric protein	Salzman et al. (2011)
	Synovial sarcoma	SS18-SSX1	70	Interchromosomal translocation	Oncogenic chimeric protein	Turc-Carel et al. (1987) and Clark et al. (1994)
		SS18-SSX2	30	Interchromosomal translocation	Oncogenic chimeric protein	Crew et al. (1995)
		SS18-SSX4	0–5	Interchromosomal translocation	Oncogenic chimeric protein	Skytting et al. (1999)

Please cite this article in press as: M.J. Annala et al., Fusion genes and their discovery using high throughput sequencing, Cancer Lett. (2013), http://dx.doi.org/10.1016/j.canlet.2013.01.011