

ARTICLE OPEN



Molecular Diagnostics

Genomic architecture of FGFR2 fusions in cholangiocarcinoma and its implication for molecular testing

Olaf Neumann^{1,2,3}, Timothy C. Burn^{4,14}, Michael Allgäuer¹, Markus Ball^{1,5}, Martina Kirchner¹, Thomas Albrecht¹, Anna-Lena Volckmar¹, Susanne Beck¹, Volker Endris¹, Hannah Goldschmid¹, Ulrich Lehmann⁶, Huriye Seker-Cin¹, Sebastian Uhrig^{7,8}, Stephanie Roessler¹, Jan Budczies^{1,2}, Stefan Fröhling^{2,3,8,9}, Thomas Longerich^{1,10}, Alex H. Wagner^{11,12}, Arndt Vogel¹³, Peter Schirmacher^{1,2,3,10}, Albrecht Stenzinger^{1,2,3,5} and Daniel Kazdal^{1,2,3,5}

© The Author(s) 2022

BACKGROUND: Cholangiocarcinoma (CCA) is a primary malignancy of the biliary tract with a dismal prognosis. Recently, several actionable genetic aberrations were identified with significant enrichment in intrahepatic CCA, including *FGFR2* gene fusions with a prevalence of 10–15%. Recent clinical data demonstrate that these fusions are druggable in a second-line setting in advanced/metastatic disease and the efficacy in earlier lines of therapy is being evaluated in ongoing clinical trials. This scenario warrants standardised molecular profiling of these tumours.

METHODS: A detailed analysis of the original genetic data from the FIGHT-202 trial, on which the approval of Pemigatinib was based, was conducted.

RESULTS: Comparing different detection approaches and displaying representative cases, we described the genetic landscape and architecture of *FGFR2* fusions in iCCA and show biological and technical aspects to be considered for their detection. We elaborated parameters, including a suggestion for annotation, that should be stated in a molecular diagnostic *FGFR2* report to allow a complete understanding of the analysis performed and the information provided.

CONCLUSION: This study provides a detailed presentation and dissection of the technical and biological aspects regarding *FGFR2* fusion detection, which aims to support molecular pathologists, pathologists and clinicians in diagnostics, reporting of the results and decision-making.

British Journal of Cancer; <https://doi.org/10.1038/s41416-022-01908-1>

INTRODUCTION

Histological, molecular and therapeutic aspects of cholangiocarcinoma

Cholangiocarcinoma (CCA), representing the second most frequent type of primary liver cancer, is a rare cancer with an annual incidence ranging from 0.72 to 1.62 per 100,000 individuals in the United States [1]. CCA exhibits an aggressive course, which is further aggravated by late presentation in most patients, resulting in a dismal prognosis with a 5-year survival rate of 5–20% [2].

According to its anatomical location, CCA is categorised into intra- and extrahepatic cholangiocarcinoma (iCCA and eCCA) with the latter being further subdivided into perihilar (pCCA) and distal cholangiocarcinoma (dCCA). Though clinical management is still largely based on this classification scheme, recent studies provide

accumulating evidence that for iCCA this classification is not sufficiently detailed [3–5]. Considering shared clinical, aetiological, histological and molecular features, as described by the new WHO classification of gastrointestinal tumours, iCCA is now further subclassified into a small- and large-duct type [6] (Suppl. 1). While large-duct type iCCA is often paralleled by the presence of biliary precursor lesions and mutations in *KRAS* (hgnc:6407) and *TP53* (hgnc:11998) similar to eCCA, genetic alterations in *IDH1/2* (hgnc:5382/hgnc:5383), *BAP1* (hgnc:950), *BRAF* (hgnc:1097) and *FGFR2* (hgnc:3689) are almost exclusively found in small-duct type iCCA [4, 7]. Given the potential druggability of the latter genetic events, this subdivision will gain increasing relevance with the rise of targeted molecular assays for the clinical management of unresectable cases.

¹Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany. ²Center for Personalized Medicine Heidelberg (ZPM), Heidelberg, Germany. ³German Cancer Consortium (DKTK), Heidelberg, Germany. ⁴Formerly employed by Incyte Research Institute, Wilmington, DE, USA. ⁵Translational Lung Research Center (TLRC) Heidelberg, German Center for Lung Research (DZL), Heidelberg, Germany. ⁶Institute of Pathology, Hannover Medical School, Hannover, Germany. ⁷Computational Oncology Group, Molecular Precision Oncology Program, National Center for Tumor Diseases (NCT) Heidelberg, Heidelberg, Germany. ⁸German Cancer Research Center (DKFZ), Heidelberg, Germany. ⁹Department of Translational Medical Oncology, National Center for Tumor Diseases (NCT) Heidelberg, Heidelberg, Germany. ¹⁰Liver Cancer Center Heidelberg, Heidelberg, Germany. ¹¹The Steve and Cindy Rasmussen Institute for Genomic Medicine, Nationwide Children's Hospital, Columbus, OH, USA. ¹²Department of Pediatrics and Biomedical Informatics, The Ohio State University College of Medicine, Columbus, OH, USA. ¹³Gastroenterology, Hepatology, and Endocrinology, Hannover Medical School, Hannover, Germany. ¹⁴Present address: Tyra Biosciences, Carlsbad, CA, USA. ✉email: albrecht.stenzinger@med.uni-heidelberg.de; daniel.kazdal@med.uni-heidelberg.de

Received: 7 February 2022 Revised: 23 June 2022 Accepted: 29 June 2022

Published online: 23 July 2022

Studies have shown that up to 20% of iCCA harbour *IDH1* and *IDH2* mutations [8, 9]. Activating mutations and gene fusions involving *FGFR2* were mutually exclusive with these and identified in 4% [10] and 10–15% [11, 12] of intrahepatic bile duct carcinomas, respectively. *FGFR2* alterations are restricted to the small-duct subtype of iCCA. This study focuses on *FGFR2* fusions that exhibit a diverse and steadily growing number of fusion partners and translocation breakpoints. Over the last few years, selective oral inhibitors targeting *FGFR1/2/3* (hgnc:3688/3689/3690) were developed. A recent multicenter, open-label, single-arm, multicohort, Phase II study (FIGHT-202) [13], which investigated the *FGFR2*-inhibitor Pemigatinib (INCB054828), reported an overall response rate (ORR) of 35.5% and a disease control rate (DCR) of 82% in 107 patients with *FGFR2* fusions or associated rearrangements. These data led to the approval of Pemigatinib as a second-line treatment for patients with metastatic/advanced CCA harbouring a *FGFR2* fusion or another *FGFR2*-associated rearrangement, first by the FDA and more recently by the EMA. Based on these encouraging data, an open-labelled, randomised, Phase III study (FIGHT-302) is underway to evaluate Pemigatinib as first-line treatment vs. conventional chemotherapy. Moreover, the *FGFR* inhibitors Erdafitinib and Rogaratinib were evaluated in CCA patients as part of an entity-independent basket trial and Erdafitinib was tested in a Phase II study of 14 patients with biliary tract cancers and *FGFR* alterations including activating mutations and fusions. The study demonstrated a disease control rate of 83.3% with a median progression-free survival of 5.6 months. Larger Phase II studies with Infigratinib (BGJ398) [14] and Futibatinib [15] showed an objective response rate of 23.1% and 25.4% for fusion-positive iCCA, respectively. Taken together, these data herald a paradigm shift in the clinical management of iCCA, which translates compellingly into comprehensive molecular profiling of these tumours to reflect the new standard of care. It will be of importance to establish reliable molecular pathological methods in routine diagnostics to be able to identify all treatable genetic alterations especially *FGFR2* gene fusions and rearrangements. Given the immense diversity of *FGFR2* fusions, both a detailed understanding of the genetic landscape of *FGFR2* aberrations and thorough knowledge about the technology and methodology employed for detection is crucial. Here we report the detailed analysis of the diverse genetic architecture of *FGFR2* gene fusions and derive implications for selection of available methods in molecular diagnostics as well as for the annotation in diagnostic reports. We put a specific focus on the advantages and disadvantages of widely-used, focused, next-generation sequencing-based methodologies, which will to our mind become the gold standard for gene fusion analysis in a routine clinical setting.

Detailed consideration of *FGFR2*-associated rearrangements

***FGFR2* Structure: exonic regions and protein domains.** The fibroblast growth factor receptor 2 (*FGFR2*) is a member of a gene family of four receptor tyrosine kinases (*FGFR1-4*) that are involved in several crucial cell functions including angiogenesis, differentiation, development, survival, tissue repair, and proliferation [16]. The gene locus of *FGFR2* is located on the reverse strand of the long arm of chromosome 10 (10q26.13) spanning nucleotide positions 123,237,844 to 123,357,972 (hg19, alignment from ncbigene:2263). At the DNA level, at least 24 (according to NCBI; www.ncbi.nlm.nih.gov/gene/2263) regions exist that qualify as exonic sequences, but only subsets of these are used for different isoforms through alternative splicing, resulting in a vast variability on transcript level. To the present day, more than 25 isoforms of *FGFR2* have been described.

The canonical isoform of *FGFR2* refers to the transcript NM_000141.4 (also named *FGFR2* IIIc), consisting of 18 exons which code for an 821 amino acids (aa) long protein that is expressed in mesenchymal cells. It differs from the transcript

NM_022970.3 (*FGFR2* IIIb), which is predominantly expressed by epithelial cells, only by an alternative eighth exon that alters a part of the third immunoglobulin (Ig)-like domain and results in elongation of the protein by one amino acid to a total of 822 aa. Both isoforms share three extracellular Ig-like domains (Ig-like I, aa 43–115; Ig-like II, aa 172–248; Ig-like III, aa 264–359), a heparin-binding domain (aa 162–178), a transmembrane domain (aa 371–401), a tyrosine kinase domain (aa 481–757), and a proline-rich C-terminus containing multiple tyrosine phosphorylation sites, but differ in their binding affinity for different fibroblast growth factors (FGF). Of note, for mutation and annotation purposes, the canonical isoform NM_000141.4 is mostly used, although from a biological point of view, the isoform NM_022970.3 is expected to be expressed in epithelial tumours, such as iCCA. This issue can be seen as a purely formal aspect since oncogenic alterations within the alternative eighth exon are rare and oncogenic translocations do not occur there. As long as the RefSeq ID is specified with the detected alteration, the annotation is comprehensible and can be converted to another isoform if necessary. A schematic graphical representation of these two isoforms is shown in Fig. 1. In addition, the isoform NM_001144913 which encodes the oncogenic *FGFR2* IIIb C3 is provided.

Under wild-type conditions, simultaneous binding of an FGF and heparin/heparan sulfates as co-factors to *FGFR2* leads to conformational changes that induce dimerisation and subsequently trans-autophosphorylation of the receptor at its cytoplasmic part. The activated intracellular kinase domain is now able to phosphorylate downstream targets such as *FRS2*, leading to the activation of various signalling pathways, including *MAPK*, *PI3K-AKT*, and *JAK-STAT*.

Oncogenic activation of *FGFR2*. The ligand-independent activation of *FGFR2* is a known oncogenic event [17] predominantly found in intrahepatic cholangiocarcinoma with a prevalence of around 13% [11, 18]. But it is also present at low frequencies in many other tumour entities including but not limited to breast, endometrial, gastric, lung, or thyroid cancer. Two main mechanisms are described for the ligand-independent activation of *FGFR2* as a result of a genetic rearrangement, which will be discussed below: (I) the loss of the C-terminal region of *FGFR2* and (II) the gain of domains enhancing dimerisation.

C-terminal truncation: Alternative splicing of *FGFR2* can also affect the C-terminal region and result in shortened proteins, described as C2 and C3 variants [19], lacking the last 34 or 53 aa, respectively. Breast cancer cell lines expressing an isoform with a C3 terminus (e.g. NM_001144913.1 or NM_001144919.2) show potent transformation capability and ligand-independent, constitutive *FGFR2* activation, indicated by strong phosphorylation of *FRS2* (hgnc:16971) [20]. Similar results showing an enhanced transforming activity of the C3 variant are described in other studies [21]. The oncogenic effect of these C-terminal deletions, which may encompass the entire last exon (exon18, NM_000141.4), is attributed to aberrant receptor internalisation and enhanced *FRS2*-dependent signalling, caused specifically by the loss of the YLDL motif [22] (aa 769–772, NM_000141.4). The proline-rich C-terminal region acts as a binding site for the SH3 domain of the growth factor receptor-bound protein 2 (*GRB2*, hgnc:4566). *GRB2* is a critical component in the downstream signalling of almost all receptor tyrosine kinases. But it was also shown to have a role in the negative regulation of *FGFR2* by sterically hindering C-terminal phosphorylation and adapter protein recruitment [23]. Truncation of the last 10 aa of *FGFR2* can impair the binding of *GRB2* and their interaction [24]. Thus C-terminal truncation, which can be introduced by a deletion, by a splice site mutation, or by a translocation, represents an independent mechanism of *FGFR2* activation. For fusions that can be assigned to this mechanism, it does not matter whether

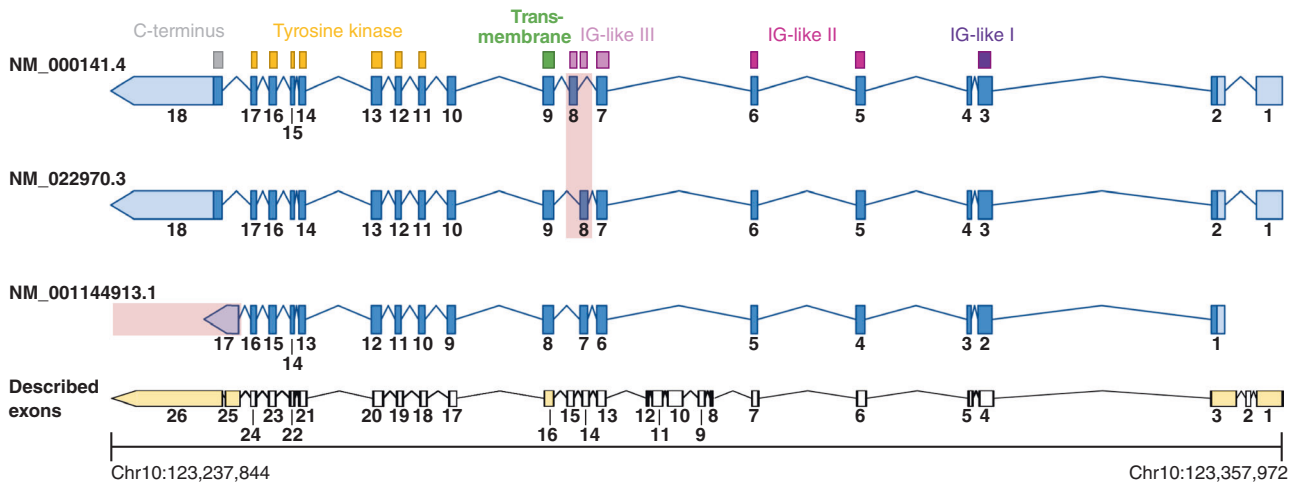


Fig. 1 Schematic representation of selected FGFR2 isoforms. Shown is the genetic structure of the two principal isoforms NM_000141.4 = FGFR2 IIIc and NM_022970.3 = FGFR2 IIIb expressed by mesenchymal and epithelial cells, respectively. Both isoforms differ only by the use of an alternative eighth exon (highlighted in red) that alters a part of the third immunoglobulin (Ig)-like domain. The isoform NM_001144913.1 represents a variant of the FGFR2 IIIb isoform lacking the first exon and using a shorter alternative last exon which leads to a shortened C-terminus also known as C3 terminus (highlighted in red). Blue bars and numbering on each isoform depict the encoded exon, where dark blue indicates coding sequences and light blue untranslated regions. At the bottom, a representation of 26 described exonic regions is displayed in black. Exons highlighted in yellow indicate that alternative exon regions exist at this genomic position. The lines connecting the exons represent the corresponding intronic regions but scaled down to 15% of the original length for illustrative purposes.

functional domains are obtained from the fusion partner, or whether the reading frame is retained within the fusion product.

Enhanced dimerisation: As mentioned above, binding of FGF leads to FGFR2 activation by inducing receptor dimerisation. A gene fusion leading to additional dimerisation or oligomerization domains may enhance the dimerisation ability of FGFR2, or may even make the fusion protein independent of FGF binding [11, 25]. This is observed in the vast majority of FGFR2 fusions with intact reading frames. Here, FGFR2 acts as 5' fusion partner, with breakpoints occurring downstream of the intracellular tyrosine kinase domain most frequently but not limited to within intron 17 (NM_000141.4). A multitude of different 3' fusion partners was described [26]. The common denominator of these is, that the sequence contributed to the fusion contains the code for strong dimerisation/oligomerization domains, like AFF3, BAG, BAR, coiled-coil, FN1, leucine zipper, LIS1, SAM, SPHF, or zinc finger [25, 27].

To accurately identify fusions that use this mechanism, it is imperative to identify the precise fusion junction in order to determine which protein domains are intact and functionally conserved in the fusion protein.

MATERIALS AND METHODS

A detailed genetic analysis of the original data from the FIGHT-202 trial (NCT02924376) [13], including DNA-seq [28] and RNA-seq [29] data was conducted using the R software (v.4.0.2; R Core Team, 2016).

For visualisation of genomic regions, the IGV browser (<https://igv.org/>) [30] and the UCSC genome browser (<http://genome.ucsc.edu>) [31] were used. For a schematic representation of gene fusion architectures, ARRIIBA [32] was used.

IMPLICATIONS FOR TESTING AND REPORTING

Test strategies

DNA- vs RNA-based NGS: sequencing capabilities, breakpoint vs fusion junction, transcription level, isoforms. Detection of FGFR2 rearrangements and fusions with DNA- and RNA-based approaches are routinely utilised by clinical labs. However, key differences in technologies must be considered when interpreting results. The main principles of different methods for the detection of FGFR2 fusions/translocations on DNA- or RNA level are depicted

in Fig. 2a. The break-apart FISH (BA) approach uses two probes near the FGFR2 locus that result in a combined fluorescence signal (yellow) for the wild-type situation, which is lost if a DNA break occurs between the two probes, resulting in separate fluorescence signals for each probe (green and red). The principle is reversed for the dual fusion probe (Dual). Here, two gene-specific probes are used rendering separate fluorescence signals (red and green) for the wild-type situation. In the event that the two probed loci are fused, the fluorescence signals overlap and result in a yellow fusion signal. While BA-FISH assays detect rearrangements in a partner-agnostic manner, they lack the granularity of sequencing-based methods and provide no information on gene partners. Dual-FISH assays on the other hand can identify one specific fusion partner, but their detection spectrum is limited to fusions involving this exact partner. Both FISH approaches cannot assess the expression or structure of a fusion protein.

A fusion partner-agnostic approach on RNA level that can be used to identify likely rearrangement events is the so-called imbalance assay. Here, the relative abundance of RNA molecules is determined by considering the 5' and 3' region of the FGFR2 mRNA separately and looking for an imbalance between the two ends of the transcript. Wild-type FGFR2 would result in similar counts for the 5' and 3' region, while a translocation involving the 3' region would lead to a decrease of respective counts.

Regarding different NGS approaches one can differentiate between "closed" and "open" assays. Amplicon-based approaches are considered closed because they can only detect gene fusions for which a corresponding primer pair is included in the panel design. The single-primer extension approach (SPE; e.g., multiplexed anchored PCR) is an "open" assay capable of detecting gene fusions regardless of the fusion partner. This is achieved by relying only on one gene-specific primer. The primer-binding site used for the fusion partner is introduced through ligation of an adaptor. For hybrid-capture (HyCa)-based assays, target enrichment before sequencing is performed by using sequence-specific hybridisation probes. Therefore, the targeted RNA molecules are captured regardless of whether the sequence continues as wild-type FGFR2 or with an alternative sequence as a result of a fusion event.

Applying the different assays can result in different and sometimes discordant results. Figure 2b displays exemplary results than can be expected for specific translocation events. Comparing the results for

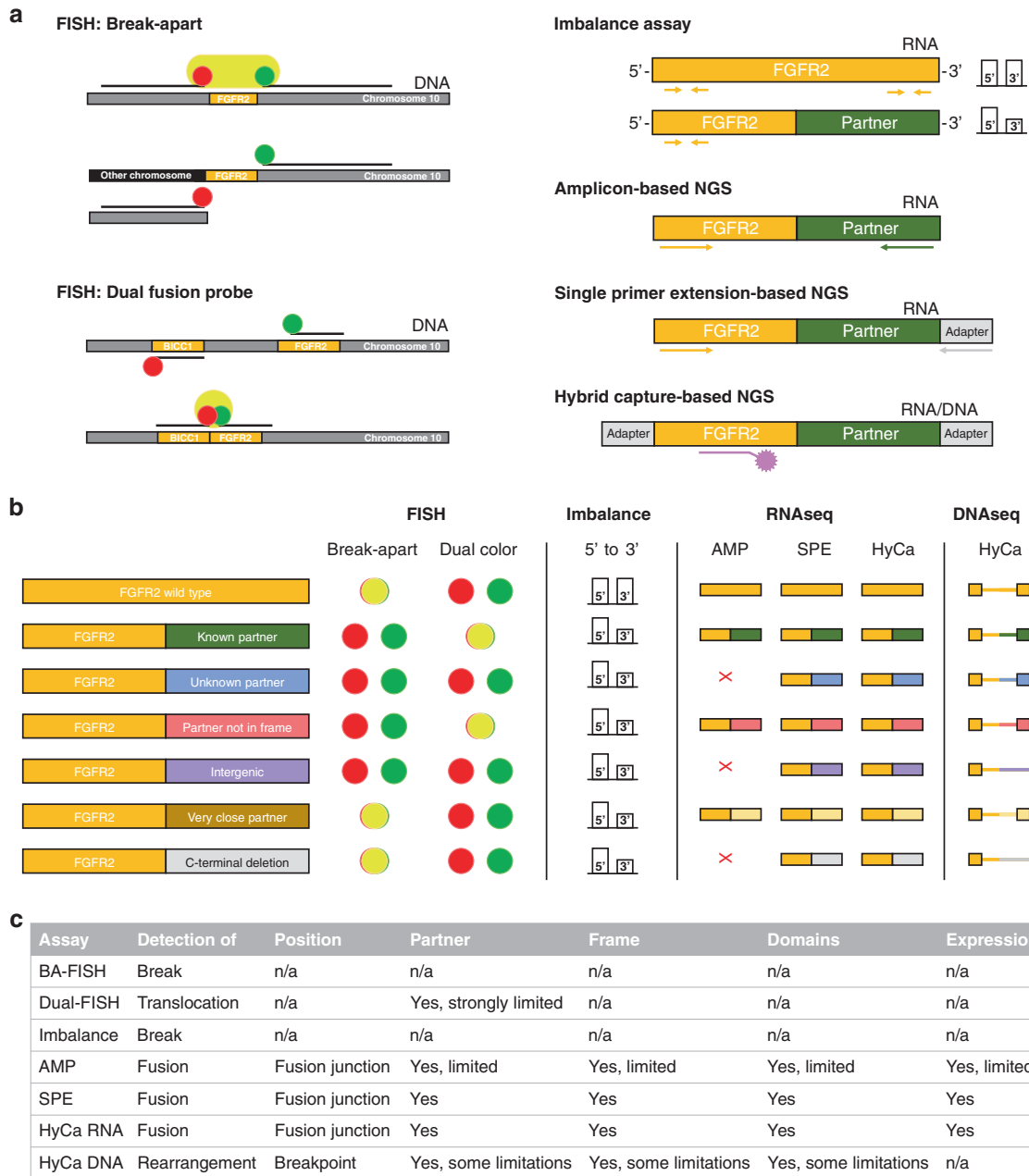


Fig. 2 Comparison of different methods for the detection of *FGFR2* translocations/fusions. **a** Main principle of fluorescence in situ hybridisation using break-apart probes (BA) and dual fusion probes (Dual), imbalance assay, amplicon (AMP)-based, single-primer extension (SPE)-based, and hybrid-capture (HyCa)-based NGS. **b** Theoretical performance of the different assays for specific fusion events. The first two rows provide the results for the wild-type and a known fusion partner as references. FISH: Separate or overlapping fluorescence signals are indicated by green and red dots or a yellow dot, respectively. Imbalance: The bars represent the number of RNA molecules detected considering the 5' or 3' region. RNA-seq and DNA-seq: Red cross indicates that a fusion event was not detectable by an assay. Wide and narrow bars represent exonic and intronic regions, respectively. **c** Characteristics and informative value of the different assays.

an unknown fusion partner, an out of frame fusion, an intergenic fusion, a very close fusion partner regarding the chromosomal position, or just a deletion of the 3' region, it becomes evident that the results of different assays provide a variable depth of detail, which can cause misinterpretation. For example, both FISH approaches and the imbalance assay are not capable of detecting whether a fusion partner is in-frame. Furthermore, 3' deletions of *FGFR2* cannot be detected by FISH and false-negative results can be expected for fusion partners that are in close proximity to *FGFR2*. While the imbalance assay would be sufficient to detect all displayed examples,

it cannot differentiate between them. In regards to NGS approaches, all examples would be detectable using SPE- or HyCa-based assays (DNA & RNA), with the added benefit that the biological underpinnings of each exemplary fusion could be identified. But the amplicon-based approach would fail to detect fusions with an unknown or intragenic partner and, if not specifically designed for it, possible 3' deletions of *FGFR2*. A summary of the details that can be gained with each assay is given in Fig. 2c.

When deciding which NGS panel to use for *FGFR2* fusion detection, there are different considerations to make. For DNA-based assays, the

actual DNA breakpoint must be sequenced, requiring tiled hybrid-capture probes across the chromosomal breakpoints, which in the case of *FGFR2* have to involve exon17, intron 17, and exon18, all downstream of the kinase domain. Chimeric sequence reads with *FGFR2* and non-contiguous genomic sequences are identified and clustered to define the breakpoint region with the gene partner being computationally inferred based on the alignment of the non-*FGFR2* sequences to the genome. As an advantage, this approach is partner gene agnostic and only relies on the alignment of the chimeric sequence cluster to the reference genome to identify partner genes. Given the computational inference, cases exist that exhibit a breakpoint in the partner gene which is not in-frame or on-strand. In these cases where a fusion partner cannot be definitively identified, the alteration is reported as a rearrangement rather than a fusion. In addition, fusions with breakpoints in the intergenic space will be reported as rearrangement with no partner gene. Published reports for *FGFR2*-rearranged cholangiocarcinoma indicate that non-partner rearrangements are reported in ~10% of all cases [28, 33]. These alterations result from a variety of molecular events including bridged fusion where non-contiguous DNA is inserted in the breakpoint between *FGFR2* and the partner gene, or true intergenic rearrangements that provide a downstream cryptic C-terminal exon. In clinical trials so far no obvious differences in response rates were reported for *FGFR2* fusions or *FGFR2* rearrangements without a partner gene identified [28]. However, further analyses are needed to investigate potential prognostic and/or predictive effects of different fusion partners, particularly with regard to the introduction of oligomerization domains or the development of resistance mechanisms.

In contrast, RNA-based methodologies for *FGFR2* fusion detection interrogate the actual fusion transcript. As such, no computational inference is needed as the exons at the fusion junction and the conservation of the reading frame are directly assessed at the level of the primary sequence, making the reliable identification of the fusion partner and the affected exons more robust and straightforward in practice. Especially for the latter, DNA-based approaches were described as an unreliable predictor [34]. But there are key assay features for evaluating an RNA-based *FGFR2* fusion detection. Most importantly the assay must be fusion partner-agnostic and should not rely on gene-specific primer pairs. This implies that amplicon-based panels are not suitable for this purpose, because they are limited to a predefined set of fusions determined by the primer design. This is best illustrated by the fact that more than 140 partner genes were so far identified for *FGFR2* fusions in CCA [33]. RNA-based, partner-agnostic assays can be based on standard RNA-seq, single-primer extension, or 5'/3' imbalance assays looking at the discordance between transcript counts for the 5' and 3' ends of the *FGFR2* mRNA. By analysing information at the transcription level, RNA-based assays provide a proof of the first step of fusion gene expression. Transcription also confers greater sensitivity to RNA-based assays compared with DNA-based assays, as it leads to a signal enhancement due to larger numbers of RNA molecules compared to DNA molecules. A disadvantage of RNA-based analyses is the instability of RNA that may especially become evident with routinely used formalin-fixed paraffin-embedded (FFPE) tissues and associated processing and physical storage conditions, that have a great impact on RNA preservation and more importantly RNA degradation.

Another important factor regarding the appropriate *FGFR2* fusion detection method is the amount of tumour cells and thereby tumour DNA/RNA in the sample. Sufficient tumour DNA/RNA is crucial for a valid identification of the alterations and to prevent false-negative results, e.g. due to low tumour cell content. An estimation of different detection limits and quality parameters is published for various assays [35, 36]. According to our (single centre) experience, small biopsies of cholangiocarcinoma are well suited for combinatorial DNA and RNA profiling by HyCa or SPE-based assays with a dropout rate of <8%. A partner-agnostic NGS approach is the best choice for detecting *FGFR2* fusions, since partner-specific approaches will perform poorly due to

the already discussed pronounced heterogeneity of fusion partners. Ideally, a combined RNA-based fusion and DNA-based translocation detection is applied. While this approach would comprehensively investigate fusion events, it may not be sufficiently efficient in terms of time, cost, and tissue consumption in a diagnostic setting. DNA translocation analysis can be pursued alternatively, if the hotspot translocation areas are sufficiently covered by primers/probes. If a sample is not sufficient for NGS-based analysis, a *FGFR2* break-apart FISH should be used for analysis, since this analysis needs only 50–100 cells for reliable detection of a possible *FGFR2* translocation. A decision tree for the selection of the appropriate technology for the detection of *FGFR2* translocations/fusions is given in Suppl. 2.

For DNA- and RNA-based assays the location of the breakpoint or fusion junction is important. Exon17 of *FGFR2* encodes the C-terminal end of the kinase domain; thus, only rearrangements involving exon17, intron 17, or the protein-coding region of exon18 will maintain the kinase domain and can act oncogenic. Given that all *FGFR2* fusions in iCCA are "type 2" fusions with C-terminal fusion partners, any rearrangements upstream of exon17 will not contain a functional kinase domain and would be expected to be insensitive to an FGFR inhibitor. As such, diagnostic reporting should reflect this consideration. A general restriction of DNA and RNA-based NGS assays is the maximum read length with high data quality between ~100 bp and 250 bp, depending on the sequencing technology. Since RNA or DNA extracted from FFPE material is degraded, the short read length is not a compromise but in fact well suited for sequencing in a routine molecular pathology setting. Sequence details covering the breakpoint or fusion junction can be obtained with a certain amount of sequence information from both partners, but this approach cannot provide information about a full-length fusion transcript.

Implications for DNA-seq-based FGFR2 translocation analysis. A direct comparison of data from samples assayed by both DNA-based NGS and RNA-seq highlights the differences between the two assay formats and their impact on fusion reporting. RNA-seq analysis was performed on two samples from the Phase 2 FIGHT-202 study (NCT02924376), which had been reported as rearrangements with no identifiable gene partner by a DNA-based NGS (Foundation Medicine). In both cases, RNA-seq revealed in-frame *FGFR2::BICC1* (hgnc:19351) fusions. Based on the RNA and DNA analysis it is clear that in these cases intergenic sequences from chromosome 1 and chromosome 10, respectively, were incorporated into the genomic breakpoint generating "bridged" fusions [29] (Fig. 3). *FGFR2* intron 17 needs to be covered in its entire length by a sufficient amount of capture probes to enable the detection of intronic breakpoints. The example shows that cases exist in which only the breakpoint in *FGFR2* could be detected by the DNA-based approach, but the complete translocation event could not be resolved, because the complex recombination involved more than two chromosome arms/areas. In these cases, RNA analysis allows direct detection of the fusion event and provides information on whether the fusion is (1) in-frame, (2) contains intact structures like kinase domains that are important for therapy decision-making and (3) is expressed at all. Examples like this highlight the importance of reporting rearrangements involving the *FGFR2* hotspot region with no definitive partner gene identified by a DNA-based assay.

Implications for FISH-based FGFR2 breakage/translocation detection. FISH-based detection of chromosomal rearrangements is a well-established molecular diagnostic technology that is also fast, reliable, and cost-efficient. Two FISH approaches widely in use are the break-apart probes (BA) and fusion-specific probes (Dual) (Fig. 2). While the first approach is mostly fusion partner-agnostic, the second is specific and limited to one exact partner gene. Additionally, specific fusion variants or the expression of a putative fusion cannot be identified with this technique. The

break-apart approach relies on the strategy that after a chromosomal rearrangement event the two probes are separated far enough from each other in order to assure visibility of the two different colour signals. Therefore, genomic rearrangements that do not sufficiently separate the two probes, as it is the case for genes that are in close proximity to each other, can yield false-negative results applying FISH analyses. In fact, this situation was shown for certain *FGFR2* rearrangements. As demonstrated in Fig. 4, a rearrangement of the *FGFR2* locus with the *ATE1* gene (hgnc:782), which is located in close proximity to *FGFR2*, cannot be detected with two commercially available probes. Similar cases, that might have been missed using FISH, were seen in 4.2% (four *TACC2* (hgnc: 11523) translocations and two intergenic translocations events in close proximity to *FGFR2*) of fusions identified in the cohort of the FIGHT-202 study.

Implications for RNA-seq *FGFR2* fusion analysis. RNA-based NGS approaches are the preferred method for the detection of oncogenic fusions. Its advantages are (I) the direct demonstration that a fusion is transcribed, indicating expression, (II) the possibility to determine if the fusion is in-frame and (III) the identification of the fusion partner and the involved exons. Practical limitations are the higher costs and the need to extract RNA from the sample. Three main techniques are currently used for the detection of *FGFR2* fusions, of which the HyCa- and SPE-based approaches are the most reliable, whereas the AMP-based technique is not preferable in this regard, given the large number and diversity of fusion partners (see above). Hybrid-capture probes cover all possible areas of the *FGFR2* gene locus that

qualify as exons (Fig. 2), while SPE-based panels are designed for the canonical *FGFR2* isoform (NM_000141.4). Table 1A gives an overview of the exonic regions covered by different panels, representative for each of the three methods. With the given translocation breakpoints of the cohort, we calculated the putative fusions that most likely will arise and summarised our findings in Table 1B and C. The SPE-based technology is able to detect up to 94–97% of the possible rearrangements/fusions, depending on the panel. The HyCa-based panels can detect all of these fusions, as they include probes targeting all regions of interest. The predicted fusion calls for the amplicon-based OCAv3-RNA Panel show that with this panel only 28.5% of the putative fusions could be detected. This is based on the fact, that for AMP-based assays to be able to detect a certain *FGFR2* fusion, a primer for the specific exon of the fusion partner involved needs to be included in the panel design. The number of all relevant exons of all possible *FGFR2* fusion partners is simply too large to be accommodated in a single AMP panel. Therefore, trade-offs must be made that may result in shortcomings in certain cases. An example for this with respect to the fusion partners detected in the cohort is *CCDC6* (hgnc:18782). While specific primers for exons 1 and 2 are included in the panel, two of the four *FGFR2*::*CCDC6* fusions queried had breakpoints in exons 3 or 4 of *CCDC6* and thus could not be detected. Closed amplicon-based assays are clearly limited in detecting of oncogenic fusions with many different partners.

The following example shows issues that may arise during annotation, especially for fusions detected on DNA level. Figure 5 displays the case of a *FGFR2* fusion with the sequence of the

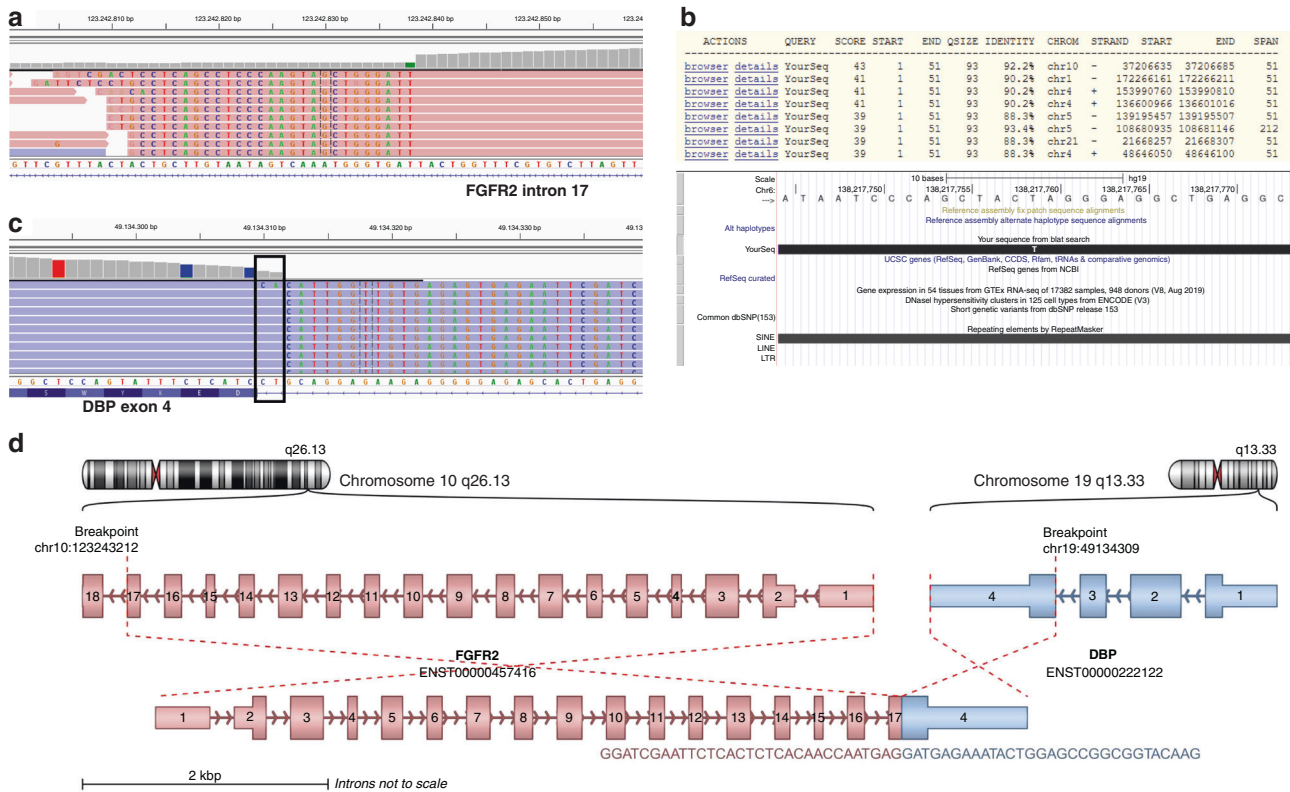


Fig. 3 Example of a *FGFR2* fusion where the fusion partner could not be identified by DNA sequencing. **a** The identified translocation point localised in intron 17 of *FGFR2* (NM_000141), shown in the IGV browser (30). The aligned soft-clipped sequences are part of an Alu repeat. **b** Upper part: Part of a BLAT result list (>100 results) of a representative split read from (a). Lower part: Expanded view of the results showing the split read sequence in reverse complement using the UCSC genome browser (31). It is part of a SINE repeat (AluSx family). **c** RNA-seq identifies the fusion partner as DBP (hgnc:2697). Displayed is the fusion junction in DBP with soft-clipped reads that align to *FGFR2*, shown in the IGV browser (30). Of note, the two marked bases belong to the *FGFR2* part (exon17) of the fusion transcript, but as they can be aligned to either fusion partner they were displayed by the IGV browser as aligned to the last two bases of DBP intron 3. **d** The putative full-length transcript of the fusion; rendered with ARRIBA (32).

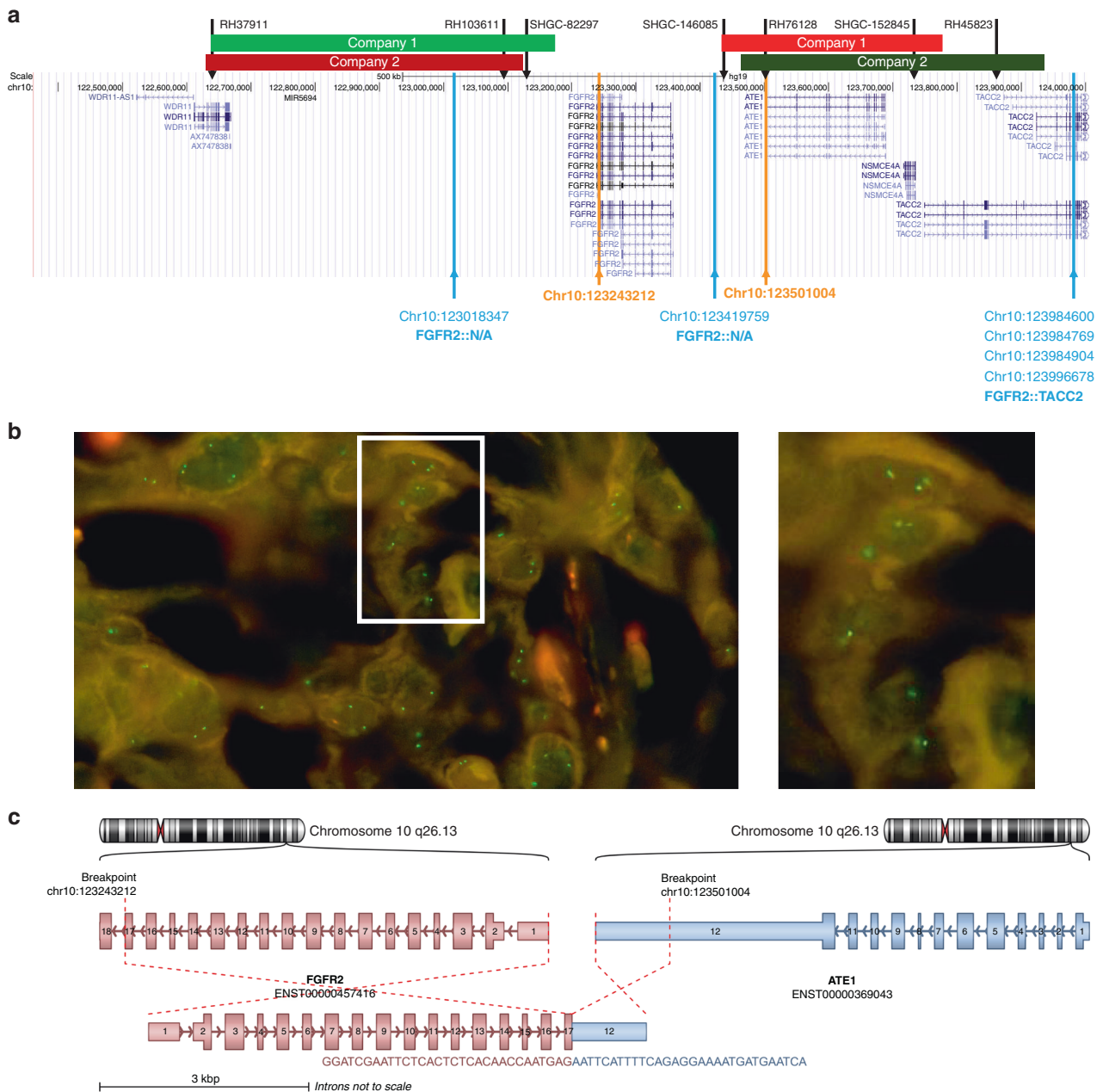


Fig. 4 Break-apart FISH analysis of a *FGFR2::ATE1* (F17A12) fusion with breakpoints after *FGFR2* exon17 and in *ATE1* exon12. **a Representation of the chromosomal region of *FGFR2* and *ATE1* using the UCSC genome browser (31). The approximate hybridisation location of the FISH probes of 2 different assays are displayed as green and red bars (light colours company 1, dark colours company 2). Orange lines show the estimated breakpoints for *FGFR2* and *ATE1*. The black arrows indicate STS markers used for mapping of the BAC clone positions. The blue arrows indicate breakpoints adjacent to *FGFR2* identified in the FIGHT-202 study. **b** Representative FISH image of the *FGFR2::ATE1* fusion-positive case. **c** Schematic representation of the putative gene fusion; rendered with ARIBBA (32).**

fusion partner aligned to *KCTD1* (hgnc:18249). As depicted, different isoforms exist for *KCTD1* with different start codons, with the isoform NM_001142730.3 representing the canonical one. In relation to the canonical isoform this fusion is not in-frame, because the translocation point lies upstream of the start codon, which classifies the *FGFR2* fusion as a C-terminal truncation (compare 1.2.2.1). But when considering other *KCTD1* isoforms (NM_001258222.3 or NM_198991.3) the breakpoint would be in intron 1 resulting in an in-frame fusion, which might lead to a fusion with enhanced dimerisation capabilities (compare 1.2.2.2). Since no RNA-based analysis was performed here, it is not possible to determine which of these hypotheses applies.

PARAMETERS TO INCLUDE IN A CLINICAL TRANSLOCATION/FUSION REPORT

Considering the technical and biological aspects of fusion analysis discussed above, we recommend to include the following information in a molecular diagnostic report concerning *FGFR2* fusions. The goal should be to provide the clinician/oncologist with a brief but concise overview of how the analysis was performed and which information is given. This includes a short description of the assays used to isolate and prepare the sample, input quantity (RNA/DNA), and the fusion detection method applied. Quality parameters like tumour cell content and specific parameters concerning the quality of the nucleic acids as well as the prepared libraries are mandatory.

Table 1. Overview of the covered FGFR2 exons by different RNA-based.

A							
FGFR2 IIIC (NM_000141.4)	Archer Dx FusionPlex Panel					Illumina	
	CTL	Expanded Lung	Lung	Oncology Research	Solid Tumor	TST170	PanCancer
Exon1	✗	✗	✗	✗	✗	✓	✓
Exon2	✓	✓	✓	✓	✓	✓	✓
Exon3	✗	✓	✗	✗	✗	✓	✓
Exon4	✗	✗	✗	✗	✗	✓	✓
Exon5	✗	✗	✗	✗	✗	✓	✓
Exon6	✗	✓	✗	✗	✗	✓	✓
Exon7	✓	✓	✓	✗	✗	✓	✓
Exon8	✓	✓	✓	✓	✓	✓	✓
Exon9	✓	✓	✓	✓	✓	✓	✓
Exon10	✓	✓	✓	✓	✓	✓	✓
Exon11	✗	✗	✗	✗	✗	✓	✓
Exon12	✗	✓	✗	✗	✗	✓	✓
Exon13	✗	✓	✗	✗	✗	✓	✓
Exon14	✗	✓	✗	✗	✗	✓	✓
Exon15	✗	✓	✗	✗	✗	✓	✓
Exon16	✗	✓	✓	✗	✗	✓	✓
Exon17	✓	✓	✓	✓	✓	✓	✓
Exon18	✗	✓	✗	✗	✗	✓	✓

B					
Method	Panel	Exon 17 (NM_000141.4)	Exon 18 NM_000141.4	Exon18 NM_001144915.1	Detectable fusions
SPE	CTL	132	0	0	132 (94.3%)
	Expanded Lung	132	4	0	136 (97.1%)
	Lung	132	0	0	132 (94.3%)
	Oncology Research	132	0	0	132 (94.3%)
	Solid Tumor	132	0	0	132 (94.3%)
HyCa	TST170	132	4	3	140 (100%)
	PanCancer	132	4	3	140 (100%)

C			
5' Fusionpartner	3' Fusionpartner	Detectable Fusion by OCAv3-RNA	
FGFR2 Exon 17	BICC1 Exon2	1	(82.1%)
	BICC1 Exon3	27	
	BICC1 Exon18	4	
	SHTN1 Exon7	0 (0%)	
	NOL4 Exon 7	2 (100%)	
	CCDC6 Exon 2	2 (50%)	
	AHCYL1 Exon 2	4 (100%)	
	Other partner	0 (0%)	
FGFR2	total	40 (28.6%)	

A: Overview of the exonic regions covered by different RNA-based NGS panels for the detection of FGFR2 fusions. Checked boxes: primers/capture probes for fusion detection present. Highlighted in yellow are the relevant exonic regions for the detection of oncogenic *FGFR2* fusions. B+C: Theoretical representation of the detectability of translocations from the FIGHT-202 trial with different SPE and HyCa approaches (B) or the OncoPrint comprehensive RNA v3 panel as example for a comprehensive AMP-based approach (C).

The overall number of detectable fusions is shown in bold type. NGS.

Up to date, there is no commonly accepted nomenclature for the diagnostic annotation of fusions and/or translocations. This is partly due to the different scientific communities involved. A technical annotation like the HGVS style complies with a short

exact description, but it is hardly suitable for human reading. We recommend using the consensus nomenclature under development [37] by the Variant Interpretation for Cancer Consortium (VICC), which captures information regarding the transcript,

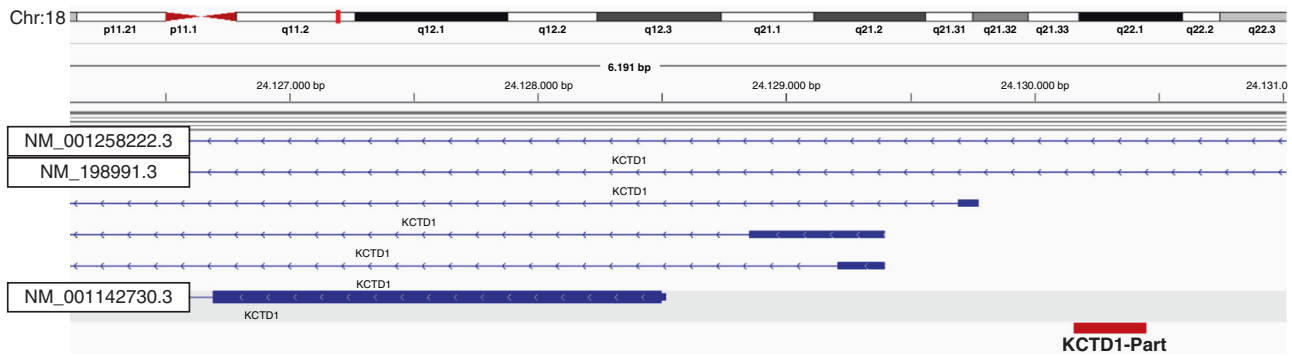


Fig. 5 Fusion annotation considering alternative isoform-specific transcription start sites. Schematic representation of the KCTD1 breakpoint of a *FGFR2::KCTD1* fusion in relation to the different isoforms encoded by this gene locus, shown in the IGV browser (30). Red bar: sequence of the fusion read aligned to KCTD1.

associated transcript exons, and the offset of the exon boundaries from the fusion junction. A first public draft of these guidelines is expected to be released in 2022.

For a diagnostic report of *FGFR2* fusions detected by RNA sequencing, we strongly recommend to include the following information (see Suppl. 3 for an example):

- The number of reads (split/paired).
- The ratio of fusion reads, as the abundance of fusions transcripts compared to wild-type transcripts.
- The gene names of both fusion partners with according transcript IDs.
- The exons involved in the fusion junction.
- The information if the fusion is in-frame and if domains crucial for an oncogenic activity are intact.

For the annotation of translocations, we endorse following the recommendations of the International System for Human Cytogenomic Nomenclature (ISCN) and the Human Genome Variation Society (HGVS) [38]. Clinical reports of translocations detected by DNA sequencing should include (see Suppl. 4 for an example):

- The number of reads (split/paired).
- The ratio of reads showing the translocation.
- The names of the involved genes.
- The chromosomal position of the identified breakpoints.
- The human genome reference build used.
- Estimations regarding the preservation of the reading frame and protein domains, should be formulated cautiously.

CONCLUSION

With the recent advancement in therapeutically targeting of *FGFR2* fusions in iCCA, standardised molecular profiling of these tumours will be necessary. Here we provide a comprehensive overview of the architecture of *FGFR2* rearrangements and discuss the implications for their reliable detection and interpretation. Comparing different approaches and based on representative cases from the FIGHT-202 trial we highlight biological and technical implications to be considered for analysis. Our data indicate that an appropriate NGS approach is recommended to reliably identify *FGFR2* gene fusion in cholangiocarcinoma. Finally, we point out parameters that should be included in a molecular diagnostic *FGFR2* report to enable full comprehension of the performed analysis and of the information provided about a detected translocation/fusion. Covering and dissecting these aspects in detail will provide a powerful resource for molecular pathologists, pathologists, and clinicians facilitating diagnostics, reporting of results as well as clinical decision-making.

DATA AVAILABILITY

Not applicable.

REFERENCES

1. Krasinakas AM. Cholangiocarcinoma. *Surg Pathol Clin*. 2018;11:403–29.
2. Yusoff AR, Razak MM, Yoong BK, Vijeyasingam R, Siti ZM. Survival analysis of cholangiocarcinoma: a 10-year experience in Malaysia. *World J Gastroenterol*. 2012;18:458–65.
3. Akita M, Sofue K, Fujikura K, Otani K, Itoh T, Ajiki T, et al. Histological and molecular characterization of intrahepatic bile duct cancers suggests an expanded definition of perihilar cholangiocarcinoma. *HPB*. 2019;21:226–34.
4. Kendall T, Verheij J, Gaudio E, Evert M, Guido M, Goeppert B, et al. Anatomical, histomorphological and molecular classification of cholangiocarcinoma. *Liver Int*. 2019;39:7–18.
5. Liao JY, Tsai JH, Yuan RH, Chang CN, Lee HJ, Jeng YM. Morphological sub-classification of intrahepatic cholangiocarcinoma: etiological, clinicopathological, and molecular features. *Mod Pathol*. 2014;27:1163–73.
6. Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, et al. The 2019 WHO classification of tumours of the digestive system. *Histopathology*. 2020;76:182–8.
7. Rizzo A, Federico AD, Ricci AD, Frega G, Palloni A, Pagani R, et al. Targeting BRAF-mutant biliary tract cancer: recent advances and future challenges. *Cancer Control*. 2020;27:1073274820983013.
8. Borger DR, Tanabe KK, Fan KC, Lopez HU, Fantin VR, Straley KS, et al. Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist*. 2012;17:72–79.
9. Kipp BR, Voss JS, Kerr SE, Barr Fritcher EG, Graham RP, Zhang L, et al. Isocitrate dehydrogenase 1 and 2 mutations in cholangiocarcinoma. *Hum Pathol*. 2012;43:1552–8.
10. Kongpetch S, Jusakul A, Lim JQ, Ng CCY, Chan JY, Rajasegaran V, et al. Lack of targetable *FGFR2* fusions in endemic fluke-associated cholangiocarcinoma. *JCO Glob Oncol*. 2020;6:628–38.
11. Arai Y, Totoki Y, Hosoda F, Shirota T, Hama N, Nakamura H, et al. Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. *Hepatology*. 2014;59:1427–34.
12. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, et al. Genomic spectra of biliary tract cancer. *Nat Genet*. 2015;47:1003–10.
13. Abou-Alfa GK, Sahai V, Hollebecque A, Vaccaro G, Melisi D, Al-Rajabi R, et al. Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *Lancet Oncol*. 2020;21:671–84.
14. Javle M, Roychowdhury S, Kelley RK, Sadeghi S, Macarulla T, Weiss KH, et al. Infigratinib (BGJ398) in previously treated patients with advanced or metastatic cholangiocarcinoma with *FGFR2* fusions or rearrangements: mature results from a multicentre, open-label, single-arm, phase 2 study. *Lancet Gastroenterol Hepatol*. 2021;6:803–15.
15. Meric-Bernstam F, Bahleda R, Hierro C, Sanson M, Bridgewater J, Arkenau HT, et al. Futibatinib, an Irreversible *FGFR1-4* inhibitor, in patients with advanced solid tumors harboring *FGF/FGFR* aberrations: a phase I dose-expansion study. *Cancer Discov*. 2021. <https://doi.org/10.1158/2159-8290.CD-21-0697>.
16. Katoh M. Fibroblast growth factor receptors as treatment targets in clinical oncology. *Nat Rev Clin Oncol*. 2019;16:105–22.

17. Neilson KM, Friesel R. Ligand-independent activation of fibroblast growth factor receptors by point mutations in the extracellular, transmembrane, and kinase domains. *J Biol Chem.* 1996;271:25049–57.
18. Graham RP, Barr Fritcher EG, Pestova E, Schulz J, Sitailo LA, Vasmataz G, et al. Fibroblast growth factor receptor 2 translocations in intrahepatic cholangiocarcinoma. *Hum Pathol.* 2014;45:1630–8.
19. Ueda T, Sasaki H, Kuwahara Y, Nezu M, Shibuya T, Sakamoto H, et al. Deletion of the carboxyl-terminal exons of K-sam/FGFR2 by short homology-mediated recombination, generating preferential expression of specific messenger RNAs. *Cancer Res.* 1999;59:6080–6.
20. Moffa AB, Ethier SP. Differential signal transduction of alternatively spliced FGFR2 variants expressed in human mammary epithelial cells. *J Cell Physiol.* 2007;210:720–31.
21. Itoh H, Hattori Y, Sakamoto H, Ishii H, Kishi T, Sasaki H, et al. Preferential alternative splicing in cancer generates a K-sam messenger RNA with higher transforming activity. *Cancer Res.* 1994;54:3237–41.
22. Cha JY, Maddileti S, Mitin N, Harden TK, Der CJ. Aberrant receptor internalization and enhanced FRS2-dependent signaling contribute to the transforming activity of the fibroblast growth factor receptor 2 IIIb C3 isoform. *J Biol Chem.* 2009;284:6227–40.
23. Lin CC, Melo FA, Ghosh R, Suen KM, Stagg LJ, Kirkpatrick J, et al. Inhibition of basal FGF receptor signaling by dimeric Grb2. *Cell.* 2012;149:1514–24.
24. Ahmed Z, George R, Lin CC, Suen KM, Levitt JA, Suhling K, et al. Direct binding of Grb2 SH3 domain to FGFR2 regulates SHP2 function. *Cell Signal.* 2010;22:23–33.
25. Wu YM, Su F, Kalyana-Sundaram S, Khazanov N, Ateeq B, Cao X, et al. Identification of targetable FGFR gene fusions in diverse cancers. *Cancer Discov.* 2013;3:636–47.
26. Li F, Peiris MN, Donoghue DJ. Functions of FGFR2 corrupted by translocations in intrahepatic cholangiocarcinoma. *Cytokine Growth Factor Rev.* 2020;52:56–67.
27. Gallo LH, Nelson KN, Meyer AN, Donoghue DJ. Functions of fibroblast growth factor receptors in cancer defined by novel translocations and mutations. *Cytokine Growth Factor Rev.* 2015;26:425–49.
28. Silverman IM, Hollebecque A, Friboulet L, Owens S, Newton RC, Zhen H, et al. Clinicogenomic analysis of FGFR2-rearranged cholangiocarcinoma identifies correlates of response and mechanisms of resistance to pemigatinib. *Cancer Discov.* 2021;11:326–39.
29. Silverman IM, Li M, Murugesan K, Krook MA, Javle MM, Kelley RK, et al. Validation and characterization of FGFR2 rearrangements in cholangiocarcinoma with comprehensive genomic profiling. *J Mol Diagn.* 2022;24:351–64. <https://doi.org/10.1016/j.jmoldx.2021.12.012>.
30. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29:24–26.
31. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. *Genome Res.* 2002;12:996–1006.
32. Uhrig S, Ellermann J, Walther T, Burkhardt P, Frohlich M, Hutter B, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res.* 2021;31:448–60.
33. Javle MM, Murugesan K, Shroff RT, Borad MJ, Abdel-Wahab R, Madison R, et al. Profiling of 3634 cholangiocarcinomas (CCA) to identify genomic alterations (GA), tumor mutational burden (TMB), and genomic loss of heterozygosity (gLOH). *J Clin Oncol.* 2019;37(15_suppl):4087.
34. Davies KD, Le AT, Sheren J, Nijmeh H, Gowan K, Jones KL, et al. Comparison of Molecular testing modalities for detection of ROS1 rearrangements in a cohort of positive patient samples. *J Thorac Oncol.* 2018;13:1474–82.
35. Kirchner M, Glade J, Lehmann U, Merkelbach-Bruse S, Hummel M, Lehmann A, et al. NTRK testing: first results of the QuiP-EQA scheme and a comprehensive map of NTRK fusion variants and their diagnostic coverage by targeted RNA-based NGS assays. *Genes Chromosomes Cancer.* 2020;59:445–53.
36. Kirchner M, Neumann O, Volckmar AL, Stogbauer F, Allgauer M, Kazdal D, et al. RNA-based detection of gene fusions in formalin-fixed and paraffin-embedded solid cancer samples. *Cancers.* 2019;11:1309.
37. Wagner AH. *Gene Fusion Curation - Sep 22, 2021*, <https://docs.google.com/presentation/d/17ggvT76n-4oc-kOjbm5rjFe1M9Pd18dwB9WY9NwKAZs/edit>.
38. Bruford EA, Antonescu CR, Carroll AJ, Chinnaiyan A, Cree IA, Cross NCP, et al. HUGO Gene Nomenclature Committee (HGNC) recommendations for the designation of gene fusions. *Leukemia.* 2021;35:3040–3.

ACKNOWLEDGEMENTS

We thank Ian Silverman for supporting the study in generating RNA-seq data.

AUTHOR CONTRIBUTIONS

Conceptualisation and design of the study: ON, TB, AS and DK. Data acquisition, analysis and interpretation: ON, TB, VE, AS and DK. Investigation: ON, TB, MA, MB, MK, A-LV, HG, HS-C, SR, JB, SF, TL, AW and AV. Writing of the original draft: ON, TB, MB, TA, AS and DK. Writing, review & editing: all authors. Visualisation: ON, JB, SB, UL, SU and DK. Supervision: PS, AS and DK. All authors agreed with the final version of the content of the manuscript.

FUNDING

This work was supported by the German Cancer Consortium (Deutsches Konsortium für Translationale Krebsforschung, DKTK) and in part by Incyte. Open Access funding enabled and organized by Projekt DEAL.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

None.

CONSENT TO PUBLISH

None.

COMPETING INTERESTS

ON reports personal fees from Novartis outside the submitted work. TB formerly employed by Incyte. AV reports personal fees from AstraZeneca outside the submitted work. JB reports a grant from the German Cancer Aid outside the submitted work. SF Consulting or advisory board membership: Bayer, Illumina, Roche; honoraria: Amgen, Eli Lilly, PharmaMar, Roche; research funding: AstraZeneca, Pfizer, PharmaMar, Roche; travel or accommodation expenses: Amgen, Eli Lilly, Illumina, PharmaMar, Roche. TL reports a personal fee from Incyte outside the submitted work. PS reports personal fees from BMS, MSD, Incyte, Janssen, Amgen, Novartis, Roche and AstraZeneca outside the submitted work. AS reports grants and personal fees from Bayer, BMS, grants from Chugai and personal fees from AstraZeneca, MSD, Takeda, Seattle Genetics, Novartis, Illumina, Thermo Fisher, Eli Lilly, Takeda, outside the submitted work. DK reports personal fees from AstraZeneca, Bristol-Myers Squibb, Pfizer, Lilly, Agilent and Takeda outside the submitted work. The remaining authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41416-022-01908-1>.

Correspondence and requests for materials should be addressed to Albrecht Stenzinger or Daniel Kazdal.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022