Junction and register choices in a computational catalogue of NR4A3 fusion gapmer designs

Author. Tristan D. McRae

Independent researcher, unaffiliated. Correspondence: trimcrae@gmail.com ORCID: 0000-0002-1823-1451

Running title. Parent pairing in NR4A3 junction gapmer designs

Abstract

A fusion-junction antisense design depends on the exact RNA join and the position of the oligonucleotide across it. We examined how these choices change predicted pairing with normal transcripts in extraskeletal myxoid chondrosarcoma. The catalogue contains 190 gapmer designs across 38 modelled NR4A3 fusion junctions, with sequence records and five complementary screens. Gapmers have a central DNA segment that can recruit an RNA-cutting enzyme; the screens measure sequence matches, not cleavage. A one-base shift of a TAF15-junction design produces an eleven-base match to normal NR4A3 through the whole DNA gap. Designs for exon-2 and exon-3 acceptors also differ, and correspondence to reported patient-derived models remains unresolved without nucleotide-level junction confirmation. At the adopted criterion of ten consecutive base pairs covering the gap, 87 of 190 designs match a mature normal parent transcript. Artificial exon joins meet the same criterion at 40.6%, versus 45.8% in the panel, so an EMC-specific excess is not established. Counts without such a match depend strongly on the cutoff. The catalogue supports explicit comparison of conditional junction and register choices, rather than selection of experimentally

established leads. Most modelled junctions are unreported in patients, and the ten-base criterion is not a measured activity threshold. No cleavage, potency, delivery, safety or therapeutic window was measured.

Keywords. antisense oligonucleotide; gapmer; RNase-H1; fusion transcript; NR4A3; extraskeletal myxoid chondrosarcoma; off-target screening

1 Introduction

Extraskeletal myxoid chondrosarcoma (EMC) is a cancer commonly driven by a gene fusion. Most EMC carries an in-frame EWSR1::NR4A3 fusion[1]: the joined RNA preserves a protein-coding reading frame. TAF15 is a substantial minority partner, and TCF12 and TFG are rare[2]. FUS, reported in two of five variant cases in a recent series[3], supplies eight modelled junctions here. Conventional cytotoxic chemotherapy has limited activity[4], though responses occur[5]. With a trialled tyrosine-kinase inhibitor, disease control exceeded response. This interpretation uses the review's response categories[4], not a figure stated in the trial report[6].

A fusion-junction gapmer could direct the RNA-cutting enzyme RNase-H1 to tumour RNA while sparing the normal parent transcripts. A gapmer contains a central DNA segment, called the gap, flanked by chemically modified regions called wings. The junction is the exact point at which RNA from the two genes joins. Junction-directed nucleic-acid agents have been reported against at least six fusion oncogenes[7-12]. Two were antisense oligonucleotides; the remainder were RNA-interference agents, including one expressed from a lentiviral vector. None against an NR4A3 fusion was found in the literature retrieved.

We ask how the exact RNA junction and the oligonucleotide register change the sequence matches encountered when designing a junction-directed gapmer. Each parent contributes part of the joined sequence, so normal-parent pairing is expected and must be described. Partial matches can fall outside a specified near-match search while still spanning the DNA gap. Comparing these definitions is useful for inspecting a design; it does not establish that one screen predicts biological activity better than another.

Two different lengths must be kept separate. RNase-H1's reported minimum DNA gap is six nucleotides, with seven to ten the working range[13]. Our criterion concerns the length of the paired duplex: ten consecutive base pairs covering the whole six-nucleotide gap. Ten is an adopted hybridisation criterion, not a measured cleavage threshold or a gap-length requirement. We also adopt, without establishing, two premises: incomplete overall pairing can permit cleavage, and mismatches in the wings are less protective than mismatches in the gap.

This study covers only the in-silico half of step one in a five-step, 2025 industry off-target framework[14]. We did not perform transcriptomics, which would measure changes across RNA transcripts. The Discussion proposes a parent-selectivity measurement related to step three.

2 Materials and methods

We used public data for all analyses and performed no laboratory work. Complete parameters, per-design tables and claim bounds are archived. Canonical transcripts of five partners and NR4A3 were obtained from Ensembl[15]. We designed 16-nucleotide molecules (16-mers) in 5-6-5 β-D-oxy-locked-nucleic-acid/DNA/β-D-oxy-locked-nucleic-acid geometry[13]. Each molecule has a six-nucleotide DNA gap between two

five-nucleotide wings. We moved this design across each junction in five positions, or registers, keeping the breakpoint inside the DNA gap. Six is the cited minimum, not the preferred gap length. The genome-wide arm is available only for 16-mers. The gap-level margin counts junction-unique gap bases on the shorter side of the breakpoint.

We applied five complementary sequence screens. They examine human RefSeq RNA alignments; an exhaustive transcript substitution search within one mismatch; unspliced parent RNA; the longest contiguous duplex through the gap in six mature wild-type parents; and unambiguous GRCh38 genome sequence, including mitochondrial sequence. Wild-type means the normal, unfused parent. Unspliced RNA still contains introns; mature RNA has had those introns removed. This adopted scope covers mature, precursor, exon-exon, non-coding and mitochondrial sequence.

We use two distinct definitions of a match. A near-match pairs at least 14 of 16 positions. Parent liability requires at least ten consecutive base pairs covering the whole gap. Here, liability means predicted sequence pairing, not demonstrated damage. We compared the panel with ten identically screened null ensembles: four shuffles, two base-frequency draws (uniform or composition-matched), and four real-parent chimeras. Two of the chimera ensembles join sequences at exon termini.

We re-scored alignments with a nearest-neighbour model of the longest consecutive paired sequence and report only energy separations. This model estimates duplex stability from adjacent base pairs. Melting-temperature calculations assume unmodified DNA:RNA at 250 nM strand[16], not locked phosphorothioate chemistry. Accordingly, we report no absolute melting temperature for the proposed reagents. Table 1 gives the unmodified-hybrid model's fusion-versus-parent difference. LNA and phosphorothioate

effects are unmodelled; this difference is not a validated bound for the modified chemistry.

3 Results

3.1 Designs at reported junctions

We selected two designs using the most prevalent junctions with published exon-resolved breakpoints. The highest-margin designs are 5′-GGGCATATCATCAAAC-3′ for EWSR1 exon 12 and 5′-GGGCATATCTTGTGTG-3′ for TAF15 exon 6, both joined to NR4A3 exon 3 (Table 1). Exon indices count from the transcript 5′ end, including non-coding exons. Counting only coding exons can select a different reagent. Both designs have gap-level margin three.

The two selected designs illustrate why the cutoff matters. Their longest wild-type duplexes covering the whole gap are eight and nine base pairs, respectively, both against TFG. Both therefore count as liable at an eight-base-pair cutoff. The TAF15 reagent remains liable at nine, and neither is liable at ten. Both also pair part of the gap at the wild-type NR4A3 exon-2/exon-3 seam. These partial duplexes are unmeasured and are not counted.

Moving a design by one base changes its parent-match classification. The sequence 5′-AGGGCATATCTTGTGT-3′ sits one register from the TAF15 reagent. It pairs 11 consecutive base pairs of wild-type NR4A3 through the whole gap. The two registers therefore have different sequence-match records and require separate assessment; these calculations do not establish a difference in cleavage or selectivity.

A search deeper than the default found many more matching transcript records than distinct gene locations. The EWSR1 and TAF15 reagents had 123 and eight sense-

strand transcript near-matches pairing the gap, respectively, but six and five gene loci. Most of the 123 are predicted transcript models. The EWSR1 reagent also matches wild-type TAF15 precursor RNA across an intron-exon boundary with two mismatches, one in the gap. Ten shared donor bases let this reagent span EWSR1, TAF15 and FUS breakpoints (Figure 1). The TAF15 reagent has no sense-strand precursor site.

The genome screen found similar total numbers of potentially hybridisable, gap-paired sites: 156 for the EWSR1 reagent and 135 for the TAF15 reagent. The EWSR1 reagent has one exact 16-base genomic match, intronic in annotated long non-coding RNA (lncRNA) ENSG00000304430. The TAF15 reagent has none. These are sequence predictions, not measured activity. Much genomic sequence is untranscribed, and transcription and cleavage at these sites are unmeasured.

The proposed chemistry adds another uncertainty. Both reagents have a phosphorothioate backbone throughout and five contiguous β-D-oxy-locked residues per wing, exceeding the two to four taken here as usual. Their high affinity may permit more mismatches than the near-match screens' two-mismatch ceiling. Both begin with a locked 5′-GGG tract. Increased sequence-dependent hepatotoxicity with high affinity is an adopted, untested premise here, not a retrieved finding. Neither reagent's potency or safety is established.

The two selected junctions have a modelled coverage of 68.4% of molecularly confirmed cases. This estimate combines the breakpoint distribution of an 18-case series[17] with a 58-case cohort[18]; it does not come from patient screening. Its 39.9%–82.8% range is not a confidence interval. Two of four inputs are fixed, no nominal level is assigned, and the calculation transfers a breakpoint distribution

between cohorts collected 21 years apart. The TAF15 arm uses three of three reported breakpoints, an upper bound.

Adding a third junction would raise modelled coverage to 79.0%. Its top-margin design is 5′-GGGCATATCTCCACGG-3′ for EWSR1 exon-13/NR4A3 exon-3. Selection names only the first two junctions by reported prevalence.

3.2 Selection and parent liabilities across 190 designs

Predicted pairing with normal parents was common across the panel. Across 38 in-frame junctions of five partners, 87 of 190 designs let a mature wild-type parent pair the whole gap over at least ten consecutive base pairs. NR4A3 forms the longest duplex for 61 designs; 85 pair one of their own two parents. Only seven of the 87 reach the conventional 14-of-16 near-match threshold, and that screen excludes parent records by name. The comparison involves different match definitions and search scope. It is not a sensitivity comparison against measured off-target activity or evidence of superiority over a comprehensive off-target assessment.

Precursor RNA adds an overlapping set of possible matches. Nineteen designs have sense-strand precursor near-matches pairing the whole gap. Counting each design once, the mature-parent and precursor union contains 93 designs, not the sum of the two counts. Gap mismatches score zero, so these counts bound the fully paired class, not all parent liability.

An apparently clear alignment screen can be incomplete. Seven searches never returned, and censoring leaves only 47 of 183 returned searches assessable. A clean count is a floor over that subset; most designs clear at the default search depth fail a deeper search.

The parent, precursor and genome screens cover all 190 designs without failures or censoring.

Some predicted off-target duplexes were nearly as stable as the intended target. Energy re-scoring found eight designs with fully paired 16-base off-target duplexes and 45 with a duplex within 2 kcal/mol of the intended target. Neither named reagent is in these classes: their nearest separations are 3.2 and 3.0 kcal/mol. These separations are upper bounds because the longest-run calculation ignores pairing on either side of a mismatch.

Longer DNA gaps reduce the fraction of designs that meet the criterion but do not remove parent liability. Counts are 87, 88 and 87 in panels of 190, 266 and 342, respectively; rates fall from 45.8% to 33.1% and 25.4%. At 5-10-5, pairing the whole gap itself requires ten base pairs. The screen misses shorter runs potentially compatible with the reported seven-to-ten working range, making 25.4% a floor.

Three designs clear every screen at the ten-base-pair cutoff. Two have no full-gap parent duplex at any length; the third pairs wild-type NR4A3 over eight bases. None targets a reported patient junction; these designs are possible controls for a future experiment, not established mechanism controls.

Choosing within each junction yields a different result from requiring every screen to be clear. It finds a design without a parent match meeting the adopted criterion for 35 of 38 junctions, including all five with published exon-resolved breakpoints. At cutoffs of nine, eight, seven and six base pairs, the counts become 31/3, 23/2, 9/0 and 6/0 (all junctions/published junctions). These cutoffs measure whole duplexes, not DNA gap length. The named TAF15 junction loses its best design at nine.

The comparison with artificial exon joins does not establish an EMC-specific excess. Exon-terminus chimeras meet the parent-liability criterion at 40.6%, versus 45.8% for the panel. The strongest null falls inside the panel's 95% interval at cutoffs seven through thirteen except eleven. The excess changes sign four times across cutoffs six through thirteen. Both the panel and the chimeras contain mostly junctions unreported in patients. The comparison therefore does not resolve a disease-specific excess over the liability generated by joining these genes' exon termini.

3.3 Computationally screened control sequences

Control sequences also need screening for parent pairing. Table 2 gives a dinucleotide-preserving scramble of each reagent: the bases are rearranged while retaining the counts of adjacent two-base sequences along the strand. Among such scrambles, 10.0% pair a parent's whole gap over at least ten bases, and 3.9% have the longest duplex with NR4A3. Passing this screen is necessary for these controls; it does not establish inertness.

4 Discussion

4.1 Junction identity determines which design is relevant

The catalogue can distinguish designs relevant to different RNA joins, but matching a design to a biological model remains unresolved. The panel joins donors to NR4A3 exon 3, its first coding exon. Two patient-derived cell models are instead reported at exon 2. No annotated NR4A3 transcript has exon 2 as its first coding exon, so annotation does not reconcile them.

The original protein-fusion filter excluded acceptors upstream of the initiation codon. That filter is inappropriate for RNase-H1 gapmers, which cleave RNA independently of reading frame. An EWSR1 exon-7/NR4A3 exon-2 fusion was sequenced in one of five EWSR1-rearranged tumours in a transcriptome series[19]. A PGR::NR4A3 case joins exon 2 to the NR4A3 5′ untranslated region[20]. All retrieved exon-resolved acceptors are exon 2, exon 3 or a cryptic exon in intron 2[21]. None is downstream of exon 3, where no patient-grounded designs are made.

Engineered constructs could test the two named junctions directly. A functional study reports the exon spans of three such constructs[21]. Two, E-N and T-N*, carry exactly the named reagents' junctions. Rebuilding them offers a direct test of junction-selective knockdown. However, overexpressing complementary DNA in a different setting cannot establish activity at the endogenous gene locus.

The only fusion-positive EMC cell source identified here consists of two patient-derived models with verified identity: USZ20-EMC1 (RRID:CVCL_C6MX) and USZ22-EMC2 (RRID:CVCL_C6MY)[22]. They are available on request, with no repository deposit and reported doubling times of five to six days. Their reported junctions join EWSR1 exon 13 and TAF15 exon 6 to NR4A3 exon 2. The report supplies no sequenced boundary, transcript accession or junction sequence. We therefore cannot determine whether the exon-2 label denotes a non-coding acceptor or unsupported numbering. An earlier version of this work was withdrawn after this class of indexing error.

The cell-model correspondence depends on how the reported exon labels are interpreted. A first-coding-exon acceptor is more parsimonious for the defining chimeric transcription factor[1,21]. A non-coding exon-2 acceptor would instead retain the NR4A3 initiation codon under the donor promoter. On the former interpretation,

USZ22-EMC2 matches the named TAF15 reagent and USZ20-EMC1 matches the third, exon-13 design. This remains inference. The builder emits exon-2 acceptors only for published patient seams or user-supplied sequencing checked against its transcript models.

Alternative exon-2 reagents are 5′-AGTGGGCTCTCCACGG-3′ (EWSR1 exon 13) and 5′-AGTGGGCTCTTGTGTG-3′ (TAF15 exon 6), both at top margin and below the ten-base-pair criterion. Their longest full-gap parent duplexes are eight bases with EWSR1 and nine with NR4A3, respectively. The latter directly involves the acceptor used in the selectivity ratio. Reagents cannot be exchanged between acceptors.

RNA sequencing must establish the test article's breakpoint at nucleotide resolution before ordering a reagent. Only nine of 176 distinct panel sequences match multiple junctions. Routine break-apart NR4A3 fluorescence in situ hybridisation detects rearrangement regardless of partner[4], but does not identify the nucleotide seam.

4.2 Conditional experimental proposal outside the computational findings

The following retained proposal illustrates measurements that could address biological selectivity. It is not a result of the present study or an optimised experimental protocol. We propose comparing fusion-positive and fusion-negative cells with matched genetic backgrounds, alongside the screened controls. Such isogenic comparisons have been used for NAB2::STAT6 in solitary fibrous tumour[23]. Knockdown alone cannot distinguish the relevant failure modes.

We define selectivity as a ratio of two knockdown IC50 values: wild-type NR4A3 divided by the fusion, measured from matched dose responses in the same wells. IC50

is the concentration giving half-maximal knockdown. The ratio and its threshold of 5.0 are adopted conventions, not specified by framework step three[14].

The proposed test's power depends on an unmeasured variance. Assuming a replicate standard deviation of 0.35 for the natural-log selectivity ratio, six independent biological replicates give about 80% power to falsify true selectivity of 3; three give about 30%.

High variability can make this falsification test uninformative. The relevant realised standard deviations are approximately 0.65 at three replicates, 1.53 at six, and 2.25 at ten. Above those values, no observed ratio at least one can put the upper limit of a two-sided 95% interval below 5. The test can then fail only with an anti-selective result and is treated as void. A normal-approximation interval would change these figures.

Apply this gate to the upper confidence bound on a pilot's standard deviation. At or above the threshold for the proposed replicate count, increase replication or do not test. A ratio measuring the acceptor alone cannot exclude other parent liabilities. Measure TFG alongside NR4A3, given its eight- and nine-base full-gap duplexes with the named reagents.

4.3 Interpretation and limits

The contribution is the traceable relationship between an assumed RNA junction, a chosen register and its sequence-match record. A one-base register change and a different acceptor can alter the record materially. This remains useful even without demonstrating an EMC-specific excess of parent pairing. The catalogue should be read as conditional models for inspecting these choices, not as a list of selectively active reagents.

A sequence can be designed for every modelled in-frame junction, but discrimination remains unproven. A design's own parents pair at most 13 bases across mature and precursor compartments. In contrast, eight off-target duplexes pair all 16 bases, five in curated records. This comparison is no liability ranking: the parent screen reads six transcripts, while the energy screen excludes parents.

The mature-parent and precursor screens also count different, overlapping classes. Mature-parent liability requires a ten-base full-gap duplex. Precursor liability requires a full-gap near-match within two mismatches.

The archive additionally screens the patient's unrearranged NR4A3 allele. Its two-mismatch ceiling gives a lower bound. No selected design fails, but two other registers at the EWSR1 exon-13/NR4A3 exon-2 seam do, showing further register dependence of the sequence-match classification.

Prior work establishes junction-directed oligonucleotides, and the determinants of gapmer off-target activity are already an active subject of study.[25] The present application contributes explicit NR4A3 junction and register comparisons, not the general discovery that partial complementarity matters. Four cited junction-specificity reports tested synthesised molecules[9-11,24]. We did not survey design pipelines and claim no priority for pre-synthesis screening. Whether sparing wild-type NR4A3 justifies a specificity cost remains unsettled: reported paralogue redundancy and dosage effects point in opposite directions.

The five screens describe possible sequence pairing, not measured duplex formation or cleavage. Their thresholds, energy approximations and incomplete transcript search outputs limit interpretation. Most modelled joins have not been reported in patients.

Separately, the prevalence calculation transfers breakpoint frequencies between cohorts and does not establish the fraction of patients with a usable or selectively active design. No potency, therapeutic window or safety has been measured, and systemic delivery to solid tumours remains unresolved. The supported use is to inspect and document a sequence-design choice after establishing the relevant RNA join.

Acknowledgments

No person other than the author contributed to this work.

CRediT authorship contribution statement

Tristan D. McRae: Conceptualization, Project administration. The author conceived and directed the project and is responsible for its content. AI assistance with analysis and manuscript preparation is disclosed below.

Statements and Declarations

Research use only, and not for administration to any person or animal. Every sequence is an unsynthesised, untested laboratory reagent. Order only from fusion-junction-aso-sequences.csv, which specifies sequence, locked residues and backbone, after RNA sequencing establishes the test article's breakpoint at nucleotide resolution.

Ethical considerations. Not applicable; no human subjects, human material or animals were involved, and no ethics approval was required.

Consent to participate. Not applicable; no participants were enrolled.

Consent for publication. Not applicable; no individual-level data are reported.

Declaration of conflicting interest. No competing financial interests exist. The author holds no patent, patent application, equity or consultancy relating to any sequence or method described here. One non-financial interest is declared: the author is a survivor of extraskeletal myxoid chondrosarcoma, the disease this work addresses.

Funding statement. No external funding; self-funded by the author.

Data availability. Code, per-design tables and screen parameters for the preceding analysis are archived at doi:10.5281/zenodo.22229096. This revision does not claim that the archived files contain its corrected interpretations. The accompanying sequence record (Supplementary File 1) and revision note (Supplementary File 2) correct the melting-temperature interpretation and qualify cell-model correspondence; sequence rows and numerical model outputs are unchanged. An earlier analysis mislocated the acceptor through coding-versus-transcript exon indexing and was withdrawn in full. Rebuilt, verified panels and the correction record are archived.

Declaration of generative AI and AI assisted technologies

Use of artificial intelligence. Claude (Anthropic) assisted with analysis code, screens, literature retrieval and checking, and manuscript drafting and review. GPT-6-Astra (OpenAI) assisted with this revision and checks against repository artefacts. Bibliographic records were retrieved from PubMed, Europe PMC or Crossref rather than generated, and citations were checked against those records. The author directed the work and is responsible for its content.

References

1. Labelle Y, J Zucman, G Stenman, LG Kindblom, J Knight, C Turc-Carel, B Dockhorn-Dworniczak, N Mandahl, C Desmaze, et al. (1995). Oncogenic conversion of a novel orphan nuclear receptor by chromosome translocation. Hum Mol Genet 4:2219–2226. doi:10.1093/hmg/4.12.2219

2. Paioli A, S Stacchiotti, D Campanacci, E Palmerini, AM Frezza, A Longhi, S Radaelli, DM Donati, G Beltrami, et al. (2021). Extraskeletal Myxoid Chondrosarcoma with Molecularly Confirmed Diagnosis: A Multicenter Retrospective Study Within the Italian Sarcoma Group. Ann Surg Oncol 28:1142–1150. doi:10.1245/s10434-020-08737-7

3. Chen X, X He, R Peng, M Chen and H Zhang. (2026). A series of extraskeletal myxoid chondrosarcomas with rare morphological and molecular variations. Histopathology 89:181–188. doi:10.1111/his.70131

4. Remiszewski P, S Falkowski, A Szumera-Ciećkiewicz, MJ Spałek, P Rutkowski and AM Czarnecka. (2025). From pathogenesis to the patient's bedside: a comprehensive review of extraskeletal myxoid chondrosarcoma. J Cancer Res Clin Oncol 151:283. doi:10.1007/s00432-025-06316-5

5. Stacchiotti S, GP Dagrada, R Sanfilippo, T Negri, I Vittimberga, S Ferrari, F Grosso, G Apice, M Tricomi, et al. (2013). Anthracycline-based chemotherapy in extraskeletal myxoid chondrosarcoma: a retrospective study. Clin Sarcoma Res 3:16. doi:10.1186/2045-3329-3-16

6. Stacchiotti S, S Ferrari, A Redondo, N Hindi, E Palmerini, MA Vaz Salgado, AM Frezza, PG Casali, A Gutierrez, et al. (2019). Pazopanib for treatment of advanced

extraskeletal myxoid chondrosarcoma: a multicentre, single-arm, phase 2 trial. Lancet Oncol 20:1252–1262. doi:10.1016/S1470-2045(19)30319-5

7. Skórski T, C Szczylik, L Malaguarnera and B Calabretta. (1991). Gene-targeted specific inhibition of chronic myeloid leukemia cell growth by BCR-ABL antisense oligodeoxynucleotides. Folia Histochem Cytobiol 29:85–89.

8. Toretsky JA, Y Connell, L Neckers and NK Bhat. (1997). Inhibition of EWS-FLI-1 fusion protein with antisense oligodeoxynucleotides. J Neurooncol 31:9–16. doi:10.1023/a:1005716926800

9. Parker Kerrigan BC, D Ledbetter, M Kronowitz, L Phillips, J Gumin, A Hossain, J Yang, M Mendt, S Singh, et al. (2020). RNAi technology targeting the FGFR3-TACC3 fusion breakpoint: an opportunity for precision medicine. Neurooncol Adv 2:vdaa132. doi:10.1093/noajnl/vdaa132

10. Ward SV, T Sternsdorf and NB Woods. (2011). Targeting expression of the leukemogenic PML-RARα fusion protein by lentiviral vector-mediated small interfering RNA results in leukemic cell differentiation and apoptosis. Hum Gene Ther 22:1593–1598. doi:10.1089/hum.2011.079

11. Shao L, I Tekedereli, J Wang, E Yuca, S Tsang, A Sood, G Lopez-Berestein, B Ozpolat and M Ittmann. (2012). Highly specific targeting of the TMPRSS2/ERG fusion gene using liposomal nanovectors. Clin Cancer Res 18:6648–6657. doi:10.1158/1078-0432.ccr-12-2715

12. Neumayer C, D Ng, D Requena, CS Jiang, A Qureshi, R Vaughan, TP Prakash, A Revenko and SM Simon. (2024). GalNAc-conjugated siRNA targeting the DNAJB1-

PRKACA fusion junction in fibrolamellar hepatocellular carcinoma. Mol Ther 32:140–151. doi:10.1016/j.ymthe.2023.11.012

13. Kauppinen S, B Vester and J Wengel. (2005). Locked nucleic acid (LNA): High affinity targeting of RNA for diagnostics and therapeutics. Drug Discov Today Technol 2:287–290. doi:10.1016/j.ddtec.2005.08.012

14. Andersson P, SA Burel, H Estrella, J Foy, PH Hagedorn, TA Harper Jr, SP Henry, JC Hoflack, EM Holgersen, et al. (2025). Assessing Hybridization-Dependent Off-Target Risk for Therapeutic Oligonucleotides: Updated Industry Recommendations. Nucleic Acid Ther 35:16–33. doi:10.1089/nat.2024.0072

15. Dyer SC, O Austine-Orimoloye, AG Azov, M Barba, I Barnes, VP Barrera-Enriquez, A Becker, R Bennett, M Beracochea, et al. (2025). Ensembl 2025. Nucleic Acids Res 53:D948–D957. doi:10.1093/nar/gkae1071

16. Sugimoto N, S Nakano, M Katoh, A Matsumura, H Nakamuta, T Ohmichi, M Yoneyama and M Sasaki. (1995). Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. Biochemistry 34:11211–11216. doi:10.1021/bi00035a029

17. Panagopoulos I, F Mertens, M Isaksson, HA Domanski, O Brosjö, S Heim, B Bjerkehagen, R Sciot, P Dal Cin, et al. (2002). Molecular genetic characterization of the EWS/CHN and RBP56/CHN fusion genes in extraskeletal myxoid chondrosarcoma. Genes Chromosomes Cancer 35:340–352. doi:10.1002/gcc.10127

18. Huang SC, JC Lee, YC Hsu, JW Tsai, YC Kao, TH Hsieh, YM Chang, KC Chang, PS Wu, et al. (2023). Extraskeletal Myxoid Chondrosarcomas: The Uncommon Clinicopathologic Manifestations and Significance of TAF15::NR4A3 Fusion. Mod Pathol 36:100161. doi:10.1016/j.modpat.2023.100161

19. Urbini M, V Indio, A Astolfi, G Tarantino, SL Renne, S Pilotti, AP Dei Tos, R Maestro, P Collini, et al. (2018). Identification of an Actionable Mutation of KIT in a Case of Extraskeletal Myxoid Chondrosarcoma. Int J Mol Sci 19:E1855. doi:10.3390/ijms19071855

20. Wilbur HC, DR Robinson, YM Wu, C Kumar-Sinha, AM Chinnaiyan and R Chugh. (2022). Identification of Novel PGR-NR4A3 Fusion in Extraskeletal Myxoid Chondrosarcoma and Resultant Patient Benefit From Tamoxifen Therapy. JCO Precis Oncol 6:e2200039. doi:10.1200/po.22.00039

21. Brenca M, S Stacchiotti, K Fassetta, M Sbaraglia, M Janjusevic, D Racanelli, M Polano, S Rossi, S Brich, et al. (2019). NR4A3 fusion proteins trigger an axon guidance switch that marks the difference between EWSR1 and TAF15 translocated extraskeletal myxoid chondrosarcomas. J Pathol 249:90–101. doi:10.1002/path.5284

22. Bangerter JL, KJ Harnisch, Y Chen, C Hagedorn, L Planas-Paz and C Pauli. (2023). Establishment, characterization and functional testing of two novel ex vivo extraskeletal myxoid chondrosarcoma (EMC) cell models. Hum Cell 36:446–455. doi:10.1007/s13577-022-00818-x

23. Li Y, JT Nguyen, M Ammanamanchi, Z Zhou, EF Harbut, JL Mondaza-Hernandez, CA Meyer, DS Moura, J Martin-Broto, HN Hayenga and L Bleris. (2023). Reduction of Tumor Growth with RNA-Targeting Treatment of the NAB2-STAT6 Fusion Transcript in Solitary Fibrous Tumor Models. Cancers (Basel) 15:3127. doi:10.3390/cancers15123127

24. Lee MS, S An, JY Song, M Sung, K Jung, ES Chang, J Choi, DY Oh, YK Jeon, et al. (2023). Cancer-Specific Sequences in the Diagnosis and Treatment of NUT Carcinoma. Cancer Res Treat 55:452–467. doi:10.4143/crt.2022.910

25. Watt AT, G Swayze, EE Swayze and SM Freier. (2020). Likelihood of Nonspecific Activity of Gapmer Antisense Oligonucleotides Is Associated with Relative Hybridization Free Energy. Nucleic Acid Ther. doi:10.1089/nat.2020.0847

Tables

All table values come from committed artifacts; none is typed by hand. Reagent values come from fusion-junction-aso-sequences.csv, the canonical machine-readable record. Control values come from aso-control-oligos.json. Values are either read directly or composed from those fields. Captions identify the source of each literature fact. Every reagent named here is a 5-6-5 phosphorothioate gapmer. If a design is advanced for future laboratory work after confirming the relevant RNA junction, its specification should be taken from the sequence file rather than transcribed from this page.

Table 1. The two illustrated designs, with their parent-duplex label. The parent-duplex column gives the longest consecutive duplex that a mature wild-type parent forms through the catalytic gap. Neither reagent reaches the ten-base-pair criterion, so the table prints the length rather than a pass mark.

The test articles are Brenca et al.'s engineered constructs (PMID: 31020999): E-N for the EWSR1 reagent and T-N* for the TAF15 reagent. Bangerter et al.'s two patient-derived models (PMID: 36316541) are reported at an NR4A3 exon-2 acceptor. Their correspondence to these reagents requires nucleotide-junction confirmation.

ΔTm is the modelled melting-temperature difference between the fusion duplex and the more stable half of the design's own target window. That parent differs from wild-type TFG, the parent found by the search reported in the duplex column. These differences use the unmodified DNA:RNA model at 250 nM strand concentration. LNA and phosphorothioate effects are unmodelled; the values are not validated predictions or bounds for the proposed modified reagents. Nothing here has been synthesised or tested, and no sequence may be administered to any person or animal.

Table 1

seam	reagent	margin	WT gap duplex (bp)	Model ΔTm (°C)
EWSR1 e12::NR4A3 e3	5′-GGGCATATCATCAAAC-3′	3	8 bp, wild-type TFG	26.6
TAF15 e6::NR4A3 e3	5′-GGGCATATCTTGTGTG-3′	3	9 bp, wild-type TFG	36.0

Table 2. The two control oligonucleotides, each screened as its reagent was. Each control is a dinucleotide-preserving scramble. It matches its reagent in length, first and last base, base composition and dinucleotide counts, but spans no junction. Each cleared the same mature-parent screen as its reagent. The Controls section explains the role of this screening step.

Table 2

control	sequence	scramble of	WT gap duplex (bp)
control-1	5′-GGGCATCAACATAATC-3′	the EWSR1 e12 :: NR4A3 e3 reagent	6 bp, wild-type TAF15
control-2	5′-GTATGTCATTGGCTGG-3′	the TAF15 e6 :: NR4A3 e3 reagent	7 bp, wild-type EWSR1

Figure legends

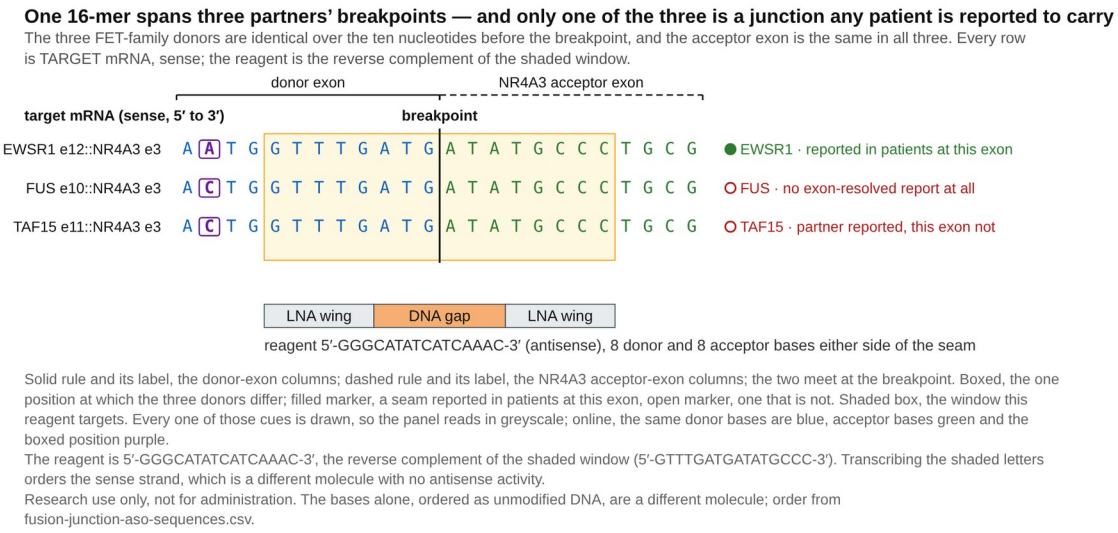

Figure 1. One 16-mer spans three partners' breakpoints; only one is reported in patients. Breakpoint-aligned windows join EWSR1 exon 12, TAF15 exon 11 and FUS exon 10 to NR4A3 exon 3, with reporting status for each row. The TAF15 exon-11 row differs from Table 1's exon-6 reagent: it is an additional breakpoint spanned by the EWSR1 reagent, not a selected reagent.

Supplementary File 2

Revision and data provenance

This note accompanies the journal-tailored manuscript “Junction and register choices in a computational catalogue of NR4A3 fusion gapmer designs”. It identifies the interpretation corrections in this submission and the relationship between its sequence record and the preceding computational archive.

Historical archive

The preceding analysis is archived at doi:10.5281/zenodo.22229096. This identifier refers to the published historical snapshot, not to the text of the present submission. The submission supplies a corrected sequence record as Supplementary File 1 and explains its interpretation here. The archive remains a source for computational code, per-design outputs and screen parameters; its superseded temperature-bound and model-correspondence statements must be read with the corrections below.

Melting temperature interpretation

The sequence record retains the legacy field names predicted_tm_fusion_c and predicted_tm_best_parent_c so its numerical outputs remain reproducible. Both are nearest-neighbour calculations for an unmodified DNA:RNA hybrid, in degrees Celsius. They are not measurements or validated predictions for the proposed gapmers, which have locked-nucleic-acid wings and a phosphorothioate backbone. Those chemical effects are not included in the model.

Subtracting the two model fields gives 26.6 and 36.0 degrees Celsius for the named EWSR1 and TAF15 reagents, respectively. The present Table 1 labels these values as model differences and removes the earlier lower-bound symbol. Neither a lower bound for the modified chemistry nor cancellation of modification effects has been established. Supplementary File 1 replaces the unsupported cancellation explanation with this limitation. Its sequence rows and numerical model outputs are unchanged from the repository source; the correction changes explanatory comments only.

Cell model correspondence

The reported NR4A3 exon-2 acceptor labels for the Zurich cell models do not by themselves establish their nucleotide junctions relative to this design panel. The present Table 1 therefore requires nucleotide-junction confirmation, consistent with the manuscript's discussion of coding-exon and transcript-exon numbering. The earlier categorical statement that these models match other designs is superseded. No cell model was obtained or sequenced in this work.

Reproduction and scope

The preserved 4 September release contains build_data.py and data-build-stamp.json. The script derives the corrected sequence record and tables from the named repository sources, checks that every non-comment CSV row is unchanged, and records SHA-256 hashes of its inputs and outputs. The upload manifest separately identifies the exact manuscript and submission files. These records document the revision; they do not replace the archived scientific computations or imply that an experiment was performed.

All sequences remain unsynthesised, untested research reagents, not for administration to any person or animal. No cleavage, potency, delivery, safety or therapeutic window was established. Experimental use requires confirmation of the test article's breakpoint at nucleotide resolution and the controls and limitations specified in the manuscript.

Journal presentation update

The Computational Biology and Chemistry package presents the completed computational work in numbered sections. The model-correspondence discussion is moved from Results to Discussion; proposed experimental evaluation remains unperformed. The title, abstract, highlights and graphical abstract summarize the same completed results. Sequence rows, numerical model outputs, table cell contents and the original scientific figure are unchanged from Qeios version 3. The historical archive limitations and the interpretation corrections described above remain in force. The current framing presents a conditional sequence-design catalogue; match classifications do not establish biological selectivity.