

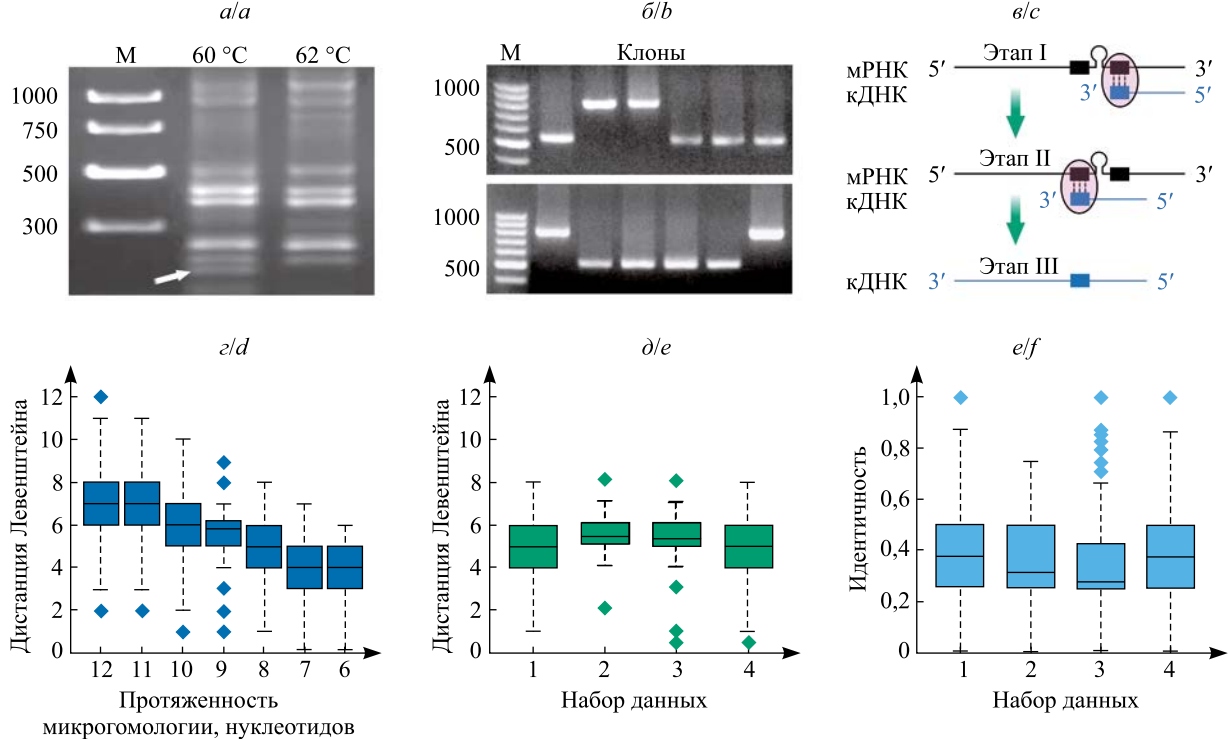


Рис. 2. Идентификация технических артефактов в первичном наборе EEJs гибридного онкогена *RUNX1-RUNXIT1* по сайтам микрогомологии:

*a* – репрезентативный ДНК гель-электрофорез ампликонов кДНК гибридного онкогена *RUNX1-RUNXIT1*, полученных при разных температурах гибридизации прямого и обратного праймеров к экзону 1 гена *RUNX1* и экзону 12а гена *RUNXIT1* соответственно. М – маркеры *Thermo Scientific™ GeneRuler™ Express DNA Ladder*, стрелкой отмечен нестабильный ампликон;

*б* – репрезентативные ДНК гель-электрофорезы результатов скрининга бактериальных клонов-трансформантов из двух независимых серий экспериментов по амплификации и клонированию фрагментов кДНК гибридного онкогена *RUNX1-RUNXIT1*. Ампликоны для клонирования были получены при использовании прямого праймера к экзону 4а гена *RUNX1* и обратного праймера к экзону 12а гена *RUNXIT1*, М – маркеры *Thermo Scientific™ GeneRuler™ 50 bp DNA Ladder*;

*в* – модель внутримолекулярной смены матрицы обратной транскриптазой.

Прямоугольниками обозначены области микрогомологии, фланкирующие стабильную вторичную структуру типа *ствол – петля* в мРНК, полупрозрачными овалами показана локализация обратной транскриптазы;

*г* – статистическая характеристика попарных дистанций Левенштейна между последовательностями длиной от 6 до 12 нуклеотидов, фланкирующими 5'- и 3'-сайты сплайсинга всех первичных EEJs;

*д* – статистическая характеристика попарных дистанций Левенштейна для 8-нуклеотидных последовательностей из разных наборов данных: последовательностей, фланкирующих 5'- и 3'-сайты сплайсинга всех первичных EEJs (1), канонических EEJs (2) и NCBI RefSeq EEJs (4), а также случайно сгенерированных 8-нуклеотидных последовательностей (3);

*е* – статистическая характеристика попарной идентичности 8-нуклеотидных последовательностей из наборов данных, описанных в части *д* этого рисунка. Для графиков *г–е* горизонтальная линия отражает значение медианы, «усы» – 1,5 межквартильного расстояния

Fig. 2. Microhomology-based identification of technical artifacts in primary set of EEJs of the fusion oncogene *RUNX1-RUNXIT1*:

*a* – representative DNA electrophoregram of amplicons of the cDNA of fusion oncogene *RUNX1-RUNXIT1*. These amplicons were generated at different annealing temperature using forward and reverse primers specific to *RUNX1* exon 1 and *RUNXIT1* exon 12a, respectively. M – markers *Thermo Scientific™ GeneRuler™ Express DNA Ladder*, arrow indicates location of unstable amplicon;

*b* – representative DNA electrophoregrams of the screening results on bacterial transformants with cloned cDNA fragments of the fusion oncogene *RUNX1-RUNXIT1*. Amplicons for cloning were generated in two independent series of experiments using forward and reverse primers specific to *RUNX1* exon 4a and *RUNXIT1* exon 12a, respectively, M – markers *Thermo Scientific™ GeneRuler™ 50 bp DNA Ladder*;

*c* – model of reverse transcriptase intramolecular template switching. Rectangles depict the microhomology sites flanking stable *stem – loop* secondary structure of mRNA, and translucent ovals show the localization of reverse transcriptase;

*d* – statistical summary on pair-wise Levenshtein distances for 6- to 12-mer oligonucleotides flanking 5' and 3' splice sites of all primary EEJs;

*e* – statistical summary on pair-wise Levenshtein distances for 8-mer oligonucleotides flanking 5' and 3' splice sites of all primary EEJs (1), canonical EEJs (2), NCBI RefSeq EEJs (4), or randomly generated 8-mer oligonucleotides (3);

*f* – statistical summary on identity of 8-mer oligonucleotides from datasets described in *e* panel.

For plots *d–f* median is shown by horizontal line, 1.5 IQR is shown by whiskers