Improved targeting and indel identification methodologies for genome editing

Lindsey A. Lonowski1, Zhang Yang1,2, Yoshiki Narimatsu, Katarzyna Duda3, Francesco Nicola1, Morten Frödi3, Henrik Clausen1,2, Hans H. Wandall1 and Eric P. Bennett1.

1 Copenhagen Center for Glycomics, Department of Cellular and Molecular Medicine and School of Dentistry, Faculty of Health Sciences, University of Copenhagen, Blegdamsvej 3, 2000 Copenhagen N, Denmark
2 Novo Nordisk Foundation Center for Bioinustainability, Danish Technical University, Lyngby, Denmark
3 Biotech Research and Innovation Centre, University of Copenhagen, Ole Maaløes Vej 5, 2000 Copenhagen N, Denmark

Conclusion
The indel detection by amplicon analysis methodology (coined IDAA18) presented here enables sensitive, precise and reliable identification of indels in a high throughput mode providing detailed information of cutting efficiency, size and nature of allelic variants generated by any of the available precise genome editing technologies (Fig.1). The IDAA1 methodology is user-friendly and easily implemented in any standard laboratory, and we expect that this screening tool will greatly advance implementation and use of precise genome targeting.

Results
Current approaches available for identification of indels include: i) enzymatic cleavage (EMC) assays14, which do not provide sensitive, reliable and absolute identification of the induced indel (Fig.2); and ii) Sanger or next generation DNA sequencing, which is costly, time and labor intensive and poorly suited for high throughput screening of hundreds or thousands of clones often required to select for desirable multi-allelic editing events that often occur at low frequency (Fig. 4).

Here, we report a novel strategy that combines use of a simple amplicon labelling strategy (Fig.1b) with the high throughput capability of DNA fragment analysis by automated Capillary Electrophoresis15 (Fig.1c) for simple detection and characterization of indels induced by genome editing technologies. The strategy is coined IDAA18 for Indel Detection by Amplicon Analysis, and we demonstrate that IDAA18 is suitable for detecting indels in both cell pools with low efficiency targeting and single sorted cells. Furthermore, we show that IDAA18 is ideally suited for high throughput detection of indels from single base events (Figs.2-3) and estimation of ‘cutting’ efficiencies of targeting tools (Fig.3).

![Figure 1](https://example.com/figure1.png)

**Figure 1. Schematic depiction of the IDAA18 strategy.** Panel a: Precise genome targeting creates double-stranded breaks that through 14-ITC introduce indels at the target site. Panel b: Tri-primer PCR at the target region amplified by use of target specific primers F71/T72 flanking the target site and a universal 5’-5mC-3’ terminal primer (Pent) specific for a 50-base sequence attached to primer T72. Tri-primer PCR results in FMD amplification labeling. Panel c: Fluorescently labeled primer containing the indels are detected by fragment analysis. Axa represent fluorescent intensity (F) and ampicron size in base pairs.

![Figure 2](https://example.com/figure2.png)

**Figure 2. Outline of the combined FACS enrichment and IDAA18 strategy.** Fluorescently labeled nucleic acids, exemplified by Cas9-2A-GFP plasmid vector (Fran et al. Nat. Protoc. 2013) with target specific gRNA (gRNA3)-induced GFP fluorescence is sorted into cell pools. 2 days post transfection cell are FACS sorted and GFP positive cells are collected for indel profiling by IDAA18.

![Figure 3](https://example.com/figure3.png)

**Figure 3. Combined FACS-mediated CRISPR/Cas9 nucleic acid targeting enrichment and IDAA18 analysis.** Panel a: GFP cell population (340 and 320) and FMDGFP (347) plasmid cut with full size and marked cosmid amplicon transfected HeLa 340 bp transfected cell population analyzed by gel electrophoresis (LSL) 326 and 283 bp for G4L and G4L2. Panel b: Lysogeny Blue 340 bp transformation of 10^6 HeLa cells analyzed by gel electrophoresis (LSL) 326 and 283 bp post transfection. Result: FACS gating strategy, P3 post GFP positive cells, P4 median GFP positive cells, P5 highest GFP positive cells. Panel c: IDAA18 analysis of unsorted and FACS enriched cell populations. Panel d: Position of the LUDOMAX.3000 standard (peaks are indicated by red line and position of wild type (WT) allele is indicated by black dotted line.

![Figure 4](https://example.com/figure4.png)

**Figure 4. Comparative EMC and IDAA analysis of ZFN targeted clones with large indels or single base indel.** EMC assays are commonly based on T4 endonuclease VII (T4E7), endonuclease V (DpnII), T7 endonuclease I (T7E1), Cell or Nanogen nucleases. Panel a: EMC (T7E1) assay of clones derived from a single LST41 clone (o10-10) targeted by Dual-GALNTZFN2. Cleaved products are indicated with asterisks. Comparative IDAA18 of the same clone shown to the right. Panel b: EMC (T7E1) assay of amplicons derived from a single HeLa clone (o10-10) targeted with COSMIC-ZFN. Comparative IDAA18 of the same clone demonstrating a monomeric TIP deletion (indicated with asterisk). Panel c: Untransfected (negative control). Unmarked minor light grey peaks represent the GSII standard.

![Figure 5](https://example.com/figure5.png)

**Figure 5. Evaluation of IDAA18 for detection of indels in gene targeted CHO cell pools and derived single clones.** Panel a: IDAA analysis of a pool of CHO cells day 2 after nucleotidization for a single base event. Panel b: Number of indels detected by MSAs deep sequencing analysis of clonal cell pools transfected with Cas9 alone, with gRNA1, gRNA4, gRNAa and gRNAy (isoform). The position of the unmodified wild type allele is indicated (a) and amplicon size (in bp) as determined by Peak Scanner software are indicated below peaks. Indel sizes determined (in bp) are shown above the most prominent peaks together with cutting efficiencies (calculated from peak area relative to total peak area in percentage). Total cutting efficiency for gRNA1 and gRNA4 were estimated to 25% and 40%, respectively, while gRNAa and gRNAy were inactive. The relative frequency of indels produced by gRNA4 was confirmed by MSAs deep sequencing panel a, which further revealed that the proportion of a deletion was an allometric insertion at a position three bases upstream of the PAM sequence (panel f). The G5UL003 standard peaks are shown in orange. Panel b: Comparative KCC analysis of the corresponding cell pools shown in (a) seven days after nucleocleotide with a monodirectional vector (STF) detecting the now revelation of truncated S-pyrimidines as a result of complete inactivation of Cas9 (arrow heads indicate single positive clones in panel c). Panel c: Representative IDAA18 analysis of single clone cell pools showing the TIP resolution power of IDAA18. Panel d: Analysis of indels by Sanger sequencing showing distribution in the 5’-5p range of individual single cell clones. Panel e: Number of indels detected by MSAs deep sequencing (Table 1) of single cell clones transfected with Cas9 alone, with gRNA1, gRNA4, gRNAa and gRNAy (isoform). Panel f: IDAA18 analysis of the corresponding cell pools transfected with Cas9 alone, with gRNA1, gRNA4, gRNAa and gRNAy (isoform). The position of the unmodified wild type allele is indicated (a) and amplicon size (in bp) as determined by Peak Scanner software are indicated below peaks. Indel sizes determined (in bp) are shown above the most prominent peaks together with cutting efficiencies (calculated from peak area relative to total peak area in percentage). Total cutting efficiency for gRNA1 and gRNA4 were estimated to 25% and 40%, respectively, while gRNAa and gRNAy were inactive. The relative frequency of indels produced by gRNA4 was confirmed by MSAs deep sequencing panel a, which further revealed that the proportion of a deletion was an allometric insertion at a position three bases upstream of the PAM sequence.