Highly Specific Enrichment of Rare Nucleic Acids using *Thermus Thermophilus* Argonaute

Jinzhao Song[1,4]*, Jorrit W. Hegge[2,4], Michael G. Mauk[1], Junman Chen[1], Neha Bhagwat[3], Jacob E. Till[3], Lotte T. Azink[2], Jing Peng[1], Moen Sen[3], Jazmine Mays[3], Erica Carpenter[3], John van der Oost[2]*, and Haim H. Bau[1]*

1. Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia PA, USA
2. Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, The Netherlands
3. Perelman School of Medicine, University of Pennsylvania, Philadelphia PA, USA

4. These authors contributed equally to this work.

*Corresponding authors: songjinz@seas.upenn.edu, john.vanderoost@wur.nl, bau@seas.upenn.edu

ABSTRACT

Characterization of disease-associated, cell-free nucleic acids (liquid biopsy) provides a powerful, minimally-invasive means for early disease detection, genotyping, and personalized therapy. Detection of alleles of clinical interest is often challenged by their low concentration and sequence homology with the much more abundant wildtype nucleic acids. Ago from the thermophilic bacterium *Thermus thermophilus* (*Tt*Ago) utilizes short DNA guides to specifically cleave complementary DNA and RNA targets. We found that under optimized conditions, *Tt*Ago cleaves DNA and RNA complementary to the guide DNA with high efficiency, but spares nucleic acids with a single nucleotide mismatch at and around its catalytic site with high sensitivity. Based on these findings, we designed a new multiplexed enrichment assay, dubbed NAVIGATER, that utilizes *Tt*Ago, to specifically cleave perfectly complementary DNA and RNA while sparing alleles of interest. NAVIGATER greatly increases the fractions of rare alleles with single nucleotide precision enhancing the sensitivity of downstream detection methods such as ddPCR, sequencing, and clamped enzymatic amplification. We demonstrate *60-fold* enrichment of *KRAS* G12D in blood samples from pancreatic cancer patients and over ten-fold improved sensitivity of clamped-PCR (PNA and XNA-PCR), enabling multiplex detection of *KRAS* and *EGFR* mutants at 0.01% fractions. *Tt*Ago has important advantages over enrichment assays such as the ones based on CRISPR-Cas. It does not require the target to contain a protospacer-adjacent motif; it is a true (turnover) catalyst; it can cleave both DNA and associated exosomal RNA targets, improving sensitivity; and can operate at elevated temperatures for higher selectivity and compatibility with detection schemes.

Introduction

In recent years, researchers have identified various enzymes from archaea and bacteria that can be programmed with nucleic acids to cleave complementary strands, among which CRISPR-Cas have attracted considerable attention as genome-editing tools[1-3]. In medical diagnostics, CRISPR- associated nucleases have been programmed to amplify reporter signals during nucleic acid detection[4-7] and enrich rare alleles to enhance detection limits of oncogenic sequences[8-10]. Despite spectacular progress, the utility of CRISPR-Cas in diagnostics is limited because it requires that the target contains a protospacer-adjacent motif (PAM)[8-10], which is absent in many sequences of clinical interest.

Analogous to CRISPR-Cas, Argonaute (Ago) proteins[11] are nucleic acid-guided endonucleases. In contrast to Cas nucleases, Ago nucleases do not require the presence of a PAM and are, therefore, more versatile than CRISPR-Cas. We found that under appropriate conditions, Ago from the thermophilic bacterium *Thermus thermophilus* (*Tt*Ago) cleaves with high efficiency DNA and RNA complimentary to the DNA guide, but spares nucleic acids with a single nucleotide mismatch at and around the catalytic site with extraordinary sensitivity. Based on these findings, we established a new PAM-independent, rare allele enrichment assay (**Fig. 1**) based on *Tt*Ago termed NAVIGATER (**N**ucleic **A**cid enrichment **Vi**a DNA **G**uided **A**rgonaute from ***T****hermus th****er****mophilus*). *Tt*Ago uses a pair of short 5′-phosphorylated single-stranded DNA oligos as guides to selectively cleave complementary double-strand (ds) alleles[12-15] while sparing alleles with a single nucleotide mismatch with guide. We optimize guide length, assay composition, and incubation temperature to enable NAVIGATER to increase rare allele fractions of *KRAS, EGFR, and BRAF* mutants in various samples. Furthermore, we demonstrate that NAVIGATER improves significantly the sensitivity of downstream detection schemes such as droplet digital PCR (ddPCR)[16], Peptide Nucleic Acid-Mediated PCR (PNA-PCR)[17], PNA-Loop Mediated Isothermal Amplification (LAMP)[18], Xenonucleic Acid clamp PCR (XNA-PCR)[19], and Sanger sequencing. As liquid biopsy (LB), especially monitoring of somatic mutations through cell-free DNA, becomes more common, methods such as NAVIGATER for enhancing sensitivity of non-invasive detection of critical diagnostic information will improve the clinical utility of such tests.

Results

Betaine, Mg^{2+}, and dNTPs enhance *Tt*Ago's cleavage efficiency of targeted nucleic acids

We envision applying NAVIGATER in combination with enzymatic amplification, in either a single-stage or a two-stage process, for rapid, inexpensive genotyping of rare mutant alleles (MAs). An enzymatic amplification process of particular interest is LAMP[20] since it does not require temperature cycling and can be implemented with simple instrumentation in resource poor settings[20]. We tested *Tt*Ago's activity in three modified variants of LAMP buffers and in a previously described custom cleaving Buffer S[15] (**Table 1).**

We incubated *Tt*Ago with either ssDNA or ssRNA fragments (100 nt) of the human *KRAS* gene and a 16 nt guide with a perfect match to the wild type (WT) *KRAS,* but with a single nucleotide mismatch at guide position 12 (g12) with *KRAS*-G12D. Cleavage products were subjected to gel electrophoresis (**Supplementary Section 1**). We define the cleavage efficiency as $\Gamma = I_C/(I_C+I_{UC})$, where I_C and I_{UC} are, respectively, band intensities of cleaved and uncleaved alleles. Comparing Γ in different buffers reveals that the cleavage efficiency in Buffer 3 is nearly 100% for both WT DNA (**Fig. 2a**) and WT RNA (**Fig. 2c**) targets, while very low (<1%) for MAs.

Next we examine why Buffer 3 outperforms the other buffers tested. Unlike the other buffers, Buffer 3 contains betaine, dNTPs, and 4×[Mg^{2+}] (8 mM vs 2 mM). To examine the effect of each of these compounds on *Tt*Ago's activity, we varied each additive's concentration in Buffers 2 and S. As the [Mg^{2+}] in Buffer S and Buffer 2 increases, so does $\Gamma_{WT\ DNA}$ at 80 °C, achieving nearly 100% at [Mg^{2+}] ~6 mM (**Fig. 2b** and **Supplementary Figs. 1a**)**.** Increase in [Mg^{2+}] has little effect on $\Gamma_{WT\ DNA}$ at 70 °C (**Supplementary Fig. 1c**). $\Gamma_{WT\ RNA}$ increases as [Mg^{2+}] increases at both 70 °C and 80 °C (**Fig. 2c** and **Supplementary Fig. 3**). Divalent cations, such as Mg^{2+} are essential for endonucleolytic activity[14,15,21,22]and stabilize the *Tt*Ago-guide-target ternary complex[20].

Betaine significantly increases $\Gamma_{WT\ DNA}$ (**Fig. 2b** and **Supplementary Figs. 1a**) but increases $\Gamma_{WT\ RNA}$ to a lesser degree **(Fig. 2d)**. Supplementing Buffer S with both Mg^{2+} (6 mM) and betaine (0.8 M) increased the *Tt*Ago cleavage efficiency from 60% to nearly 100% (**Supplementary Fig. 1a**) at 80 °C. This is consistent with betaine's ability to increase thermal stability of polymerase enzymes[23] and to dissolve secondary GC structure[24] during DNA amplification. Addition of 1.4 mM dNTPs increases $\Gamma_{WT\ DNA}$ to ~100% at 80 °C (**Supplementary Fig. 1a**) similar to betaine's effect.

At 80 °C and in the absence of dNTPs, the $\Gamma_{WT\ RNA}$ peaks at [Mg^{2+}] ~10 mM (**Supplementary Fig. 3a, left**). In contrast, in the presence of 1.4 mM dNTPs, $\Gamma_{WT\ RNA}$ increases as [Mg^{2+}] increases at both 80 °C (**Supplementary Fig. 3a, right**) and 70 °C (**Supplementary Figs. 3c** and **3d**). To ascertain that the beneficial effects of dNTPs are not unique to *KRAS*, we also tested cleavage efficiency of *EFGR* target sequences (**Supplementary Fig. 5a**). In the absence of

dNTPs, *EGFR* $\Gamma_{WT\ RNA}$ ~45% while in the presence of 1.4 mM dATP, dTTP or dCTP, $\Gamma_{WT\ RNA}$ ~100%. Surprisingly, addition of 1.4 mM dGTP does not affect cleavage efficiency. Among NTPs, only CTP increases $\Gamma_{WT\ RNA}$ (**Supplementary Fig. 5b**). Although the molecular basis of this phenomenon remains elusive, it appears that the combination of sugar groups and nitrogenous bases of dNTPs stimulates *Tt*Ago's activity.

*Tt*Ago's activity increases as pH increases from 6.6 to 9.0 (**Supplementary Fig. 5c)**. Among the buffers tested, Buffer 3 provides the best conditions for effective *Tt*Ago cleavage of targeted WT alleles likely due to the presence of betaine, dNTPs, and 8 mM [Mg^{2+}]. Notably, *Tt*Ago retains its specificity in the presence of the above additives with Γ_{MA}<1%.

Single base pair-mismatch discrimination

To cleave WT alleles efficiently while sparing alleles with single nucleotide mutations, we designed guide DNAs (gDNAs) with a single base pair mismatch with MAs. A mismatch in the seed region of mouse Argonaute (AGO2) enhances the guide-allele dissociation rate[25], reducing its cleaving efficiency. Little is known, however, on the effects of guide-allele mismatches on *Tt*Ago's catalytic activity. Molecular dynamic simulations predict that a single DNA guide - mRNA mismatch affects enzyme conformation and reduces activity, but agreement with experiments is imperfect[26]. In the absence of a reliable predictive tool, we analyze experimentally the effect on enzyme activity of a single mismatch type and mismatch position (MP) between guide and *KRAS*, *EGFR*, and *BRAF* sequences (**Fig. 3** and **Supplementary Section 2**). We use the notation MP-x to indicate that MA's aberrant nucleotide pairs with guide's x^{th} base, counted from guide's 5' end. Although all guides are complementary to their targets, variations in MP affected targets' cleavage locations (**Fig. 3a**), cleavage products' lengths (**Fig. 3b**), and, unexpectedly, cleaving efficiency (**Fig. 3c**). For example, antisense (AS) *KRAS* guide MP-5 (**Supplementary Fig. 6c)** and sense (S) *EGFR* guides MP-7 and MP-12 (**Supplementary Fig. 8b**) exhibit relatively low cleavage efficiency. This suggests that the sequence of the target-guide complex affects enzyme's conformation and activity, even when the guide and target are perfect complements.

Cleavage suppression of MAs depends sensitively on single base pair mismatch position. Mismatches in both the seed (g2-8) and mid (g9-14) regions diminished cleavage efficiency, and occasionally completely curtailed catalytic activity (**Fig. 3** and **Supplementary Section 2**). For example, *KRAS*- AS guides (15 nt) MP8-MP14 cleaved *KRAS* G12D AS (**Supplementary Fig. 6c**) and *KRAS*- S (16 nt) guides MP7 and MP11-MP13 cleaved *KRAS* G12D - S (**Fig. 3c**) with $\Gamma_{DNA\ MA}$<4%, while MP6 has $\Gamma_{DNA\ MA}$>40%. The effects of mismatch position on cleaving efficiencies of

EGFR WT and L858R are reported in **Supplementary Figs. 7** - **9**; of *EGFR* T790M in **Supplementary Figs. 10** - **11**); and of *BRAF* WT and V600E in **Supplementary Figs. 12** - **13**.

We define discrimination efficiency (DE) as the difference between *Tt*Ago-guide complex cleaving efficiency of WT and that of MA (**Fig. 3d** and **Supplementary Fig. 14**). Mismatches at and around the cleavage site (g10/g11), especially at MP7 and MP9-MP13 yielded the greatest discrimination (DE>80%) for most cases examined (**Fig. 3e**). The optimal MP depends, however, on the allele's sequence. Cleavage of RNA was more sensitive to MP than cleavage of DNA. Single mismatches at position g4-g11 nearly completely prevented RNA cleavage (**Supplementary Figs. 6e** (***KRAS* G12D**) and **9** (***EGFR* L858R**)). Our data suggests that guide's sequence differently affects the conformation of the ternary *Tt*Ago-gDNA-DNA and *Tt*Ago-gDNA-RNA complexes[21].

Short DNA guides (15/16 nt) provide best discrimination between WT and MA

*In vitro, Tt*Ago operates with ssDNA guides ranging in length from 7 to 36 nt[13]. Heterologously-expressed *Tt*Ago is typically purified with DNA guides ranging in length from 13 to 25 nt[15]. Since little is known on the effect of guide's length on *Tt*Ago*'s* discrimination efficiency (DE), we examine the effect of guide's length on DE in our in-vitro assay. *Tt*Ago efficiently cleaves WT *KRAS* with complementary guides, ranging in length from 16 to 21 nt at both 70 ^{o}C and 75 ^{o}C (**Fig 4a-i** and **Supplementary Fig. 15a-i**). Guides of 17-21 nt length with a single nucleotide mismatch MP12 cleave MAs at 75°C but not at 70^{o}C (**Fig 4a-i, top left**). Cleavage of MA at 75 ^{o}C is, however, completely suppressed with a short 16 nt guide (**Fig 4a-i**). We observe a similar behavior with guides with single mismatches at other positions (**Supplemental Fig. 16**) and with other MA sequences (**Fig 4a-ii**). Apparently, *Tt*Ago with shorter guides form a less stable complex with off-targets than longer guides thus preventing undesired cleavage.

In contrast to MA DNA, the increase in temperature did not increase undesired cleavage of MA RNA (**Figs. 4a-i, bottom and 4b-ii**), likely due to differences in the effects of ssDNA and ssRNA on enzyme conformation. When operating with a short 16 nt guide, single MP12 mismatch, and LAMP Buffer 3, *Tt*Ago efficiently cleaves both WT RNA and DNA targets while avoiding cleavage of MAs between 66 °C and 86 °C (**Fig. 4c**), providing the high specificity that is crucial for an enrichment assay.

***Tt*Ago efficiently cleaves targeted dsDNA only at temperatures above the dsDNA's melting temperature**

Guide-free *Tt*Ago can degrade dsDNA at low temperatures, and self-generate and selectively load functional DNA guides[27]. This is, however, a slow process that takes place only when target DNA is rich in AT (<17% GC)[15], suggesting that *Tt*Ago lacks helicase activity and depends on dsDNA thermal breathing to enable chopping[11,15,28]. Furthermore, since our assay is rich in gDNA that forms a tight complex with *Tt*Ago, *Tt*Ago's direct interactions with dsDNA are suppressed. *Tt*Ago*'s* ability to operate at high temperatures provides NAVIGATER with a clear advantage since dsDNA unwinds as the incubation temperature increases.

Here, we investigate *Tt*Ago's cleavage efficiency of ds*KRAS* WT and MA as functions of incubation temperature in the presence of abundant gDNA. The estimated melting temperature of 100 bp ds*KRAS* (S strand sequence listed in **Fig. 3a**) in buffer 3 is 79.7 °C[29]. Consistent with this estimate, very little cleavage takes place at temperatures under 80 ^{o}C, but *Tt*Ago cleaves dsDNA efficiently at temperatures above 80°C **(Fig. 5a)**. Target cleavage efficiency increases as the incubation time increases and saturates after about one hour (**Figs. 5c** and **5d**). Lengthier incubation time is undesirable as it leads to cleavage of MAs. For efficient discrimination between WT and MA, it is desired to incubate the assay at temperatures exceeding the target melting temperature for less than an hour.

Excess gDNA concentration is necessary to avoid off-target cleavage

At *Tt*Ago:S-guide:AS-guide ratio of 1: 0.2: 0.2, non-specific, undesired cleavage of dsMA occurs (**Fig. 5b**). This off-target cleaving becomes more pronounced as the incubation time increases (**Fig. 5b**) and can potentially be attributed to *Tt*Ago*'s* self-production of guides from dsDNA (chopping)[27]. Since gDNA forms a tight complex with *Tt*Ago, excess gDNA reduces undesired chopping. When guide concentrations exceed *Tt*Ago concentration, no apparent off-target cleavage takes place (**Figs. 5c, 5d** and **Supplementary Fig. 17**). To avoid off-target cleaving, it is necessary to saturate *Tt*Ago with guides and to limit incubation time. Indeed, at *Tt*Ago:S-guide:AS-guide ratio of 1: 10: 10, NAVIGATER efficiently cleaves double strand WT *KRAS*, *BRAF* and *EGFR* while sparing point mutations *KRAS* G12D (**Fig. 5**), *BRAF* V600E, and *EGFR* L858R (**Supplementary Fig. 18**), and *EGFR* deletion mutations in exon 19 (**Supplementary Fig. 19**).

NAVIGATER improves the sensitivity of downstream rare allele detection

In recent years, there has been a rapidly increasing interest in applying LB to detect cell-free circulating nucleic acids associated with somatic mutations for, among other things, cancer

diagnostics, tumor genotyping, and monitoring susceptibility to targeted therapies. LB is attractive since it is minimally invasive and relatively inexpensive. Detection of MAs is, however, challenging due to their very low concentrations in LB samples among the background of highly abundant WT alleles that differ from MAs by as little as a single nucleotide. To improve detection sensitivity and specificity of detecting rare alleles that contain valuable diagnostic and therapeutic clues, it is necessary to remove and/or suppress the amplification of WT alleles[30-32]. NAVIGATER meets this challenge by selectively and controllably degrading WT alleles in the sample to increase the fraction of MAs. We demonstrate here that single-plex and multiplex NAVIGATER increases sensitivity of downstream mutation detection methods such as gel electrophoresis, ddPCR[16], PNA-PCR[17], PNA-LAMP[18], XNA-PCR[19] and Sanger sequencing. Moreover, to demonstrate NAVIGATER's potential clinical utility, we enriched blood samples from pancreatic cancer patients, which have been previously analyzed with standard ddPCR protocol (**Table 2**). These samples were pre-amplified by PCR to increase WT and MA *KRAS* total content before enrichment.

Gel electrophoresis (**Supplementary Fig. 20**): We subjected enrichment assay products of pancreatic cancer patients (**Table 2**) to gel electrophoresis. In the absence of enrichment (control), the bands at 80 bp (*KRAS*) on the electropherogram are dark. After 40 minutes of *Tt*Ago enrichment, these bands faded, indicating a reduction of *KRAS* WT alleles. After 2 hours enrichment, all the bands at 80 bp, except that of patient P6, have essentially disappeared, suggesting that most WT alleles have been cleaved. The presence of an 80 bp band in the P6 lane is attributed to the relatively high (20%) MA fraction that is not susceptible to cleaving. We also PCR amplified products from a 2-hour NAVIGATER treatment, and subjected the amplicons to a second NAVIGATER (2h). The columns P3, P4, and P6 feature darker bands than P1, P2 and P5, indicating the presence of MAs in samples P3, P4, and P6 (**Supplementary Fig. 20**) and demonstrating that NAVIGATER renders observable otherwise undetectable MAs.

Droplet Digital PCR (ddPCR) (**Supplementary Fig. 21)**: To quantify our enrichment assay products, we subjected them to ddPCR. The detection limit of ddPCR is controlled by the number of amplifiable nucleic acids in the sample, which must be a small fraction of the number of ddPCR droplets. The large number of WT alleles in the sample limits the number of pre-ddPCR amplification cycles that can be carried out to increase rare alleles' concentration. Since NAVIGATER drastically reduces the number of WT alleles in the sample, it enables one to increase the number of pre-amplification cycles, increasing the number of MAs and ddPCR sensitivity. When operating with a mixture of WT and MA, NAVIGATER products include: residual

uncleaved WT (N_{WT}), MA (N_{MA}), and WT-MA hybrids (N_H). Hybrid alleles form during re-hybridization of an ssWT with an ssMA. The MA fraction is $f_{MA}=(N_{MA}+½N_H)/(N_{WT}+N_{MA}+N_H)$.

We carried out ddPCR on un-enriched (control, NAVIGATER without *Tt*Ago), once-enriched, and twice-enriched samples (**Supplementary Fig. 21b**), increasing f_{MA} significantly (**Fig. 6a**). For example, f_{MA} increased from 0.5% in the un-enriched P5 (G12D) sample to ~30% in the twice-enriched sample. This represents a ~60 fold increase in the fraction of droplets (f_{MA}) containing MA (**Fig. 6b**). The same assay also enriched G12R, increasing f_{MA} from 3% to ~66% in sample P3 and G12V, increasing f_{MA} from 5% to ~68% in sample P4 (**Fig. 6b**).

PNA-PCR: PNA-PCR engages a sequence-specific PNA blocker that binds to WT alleles, suppressing WT amplification and providing a limit of detection of f_{MA} ~1%[17]. To demonstrate NAVIGATER's utility, we compared the performance of PNA-PCR when processing pancreatic cancer patient samples (**Table 2**) before and after NAVIGATER (**Figs. 6c, 6d,** and **6e**). Before enrichment, PNA-PCR real-time amplification curves in the order of appearance are P6, P4, and P3, as expected (**Table 2**). Samples P1 (f_{MA}=0), P2 (f_{MA}=0), and P5 (f_{MA}=0.5%) nearly overlap, consistent with a detection limit of ~1%[17]. Enrichment (**Fig. 6d**) significantly increases the threshold times of samples P1 and P2, revealing the presence of MAs in sample P5. PNA-PCR combined with NAVIGATER provides the linear relationship $T_{1/2}= 22.9-5\ log(f_{MA})$ between threshold time (the time it takes the amplification curve to reach half its saturation value) and allele concentration (**Fig. 6e**), allowing one to estimate MA concentration. The data suggests that NAVIGATER can improve PCR-PNA limit of detection to below 0.1%.

PNA-LAMP: Genotyping with PNA blocking oligos can be combined with the isothermal amplification LAMP[18]. To demonstrate the feasibility of genotyping at the point of care and resource-poor settings, we use a minimally-instrumented, electricity-free Smart-Connected Cup (SCC)[20] with smartphone and bioluminescent dye-based detection to incubate PNA-LAMP and detect reaction products. To demonstrate that we can also detect RNA alleles, we used simulated samples comprised of mixtures of WT *KRAS* mRNA and *KRAS*-G12D mRNA. In the absence of pre-enrichment, SSC is unable to detect the presence of 0.1% *KRAS*-G12D mRNA whereas with pre-enrichment 0.1% *KRAS*-G12D mRNA is readily detectable (**Fig. 6f**).

Sanger Sequencing: In the absence of enrichment, Sanger sequencers detect >5% MA fraction[33]. The Sanger sequencer failed to detect the presence of f_{MA}~3% and 0.5% *KRAS-G12D* mRNA in our un-enriched samples, but readily detected these MAs following NAVIGATER enrichment (**Fig. 6g**).

XNA-PCR (**Supplementary Section 5)**: XNA-PCR[19] is a clamped assay that suppresses amplification of WT alleles, enabling detection of MAs down to 0.1% fraction. We used multiplexed NAVIGATER with 3 different guides to enrich samples of 60 ng cfDNA that included WT and various fractions of *KRAS* G12D, *EGFR* ΔE746 - A750, and *EGFR* L858R. Without NAVIGATER, XNA-PCR was able to detect down to 0.1% *KRAS* G12D (**Fig. 6h**), 0.1% *EGFR* ΔE746 - A750 (**Supplementary Fig. 23e**), and 1% *EGFR* L858R (**Supplementary Fig. 23c**). With NAVIGATER pre-treatment, XNA-PCR sensitivity increased by over 10 fold to 0.01% *KRAS* G12D (**Fig. 6i**), 0.01% *EGFR* ΔE746 - A750 (**Supplementary Fig. 23f**), and 0.1% *EGFR* L858R (**Supplementary Fig. 23d**). Here, in addition to significantly improving the sensitivity of XNA-PCR, we also demonstrate that NAVIGATER can operate as a multiplexed assay, enriching for multiple MAs.

Discussion

LB is a simple, minimally invasive, rapidly developing diagnostic method to analyze cell-free nucleic acid fragments in body fluids and obtain critical diagnostic information on patient health and disease status. Currently, LB can help personalize and monitor treatment for patients with advanced cancer, but the sensitivity of available tests is not yet sufficient for patients with early stage disease[32] or for cancer screening. Detection of alleles that contain critical clinical information is challenging since they are present at very low concentrations among abundant background of nucleic acids that differ from alleles of interest by as little as a single nucleotide.

Here, we report on a novel enrichment method (NAVIGATER) for rare alleles that uses *Tt*Ago. *Tt*Ago is programmed with short ssDNA guides to specifically cleave guide-complementary alleles and stringently discriminate against off-targets with a single nucleotide precision. Sequence mismatches between guide and off-targets reduce hybridization affinity and cleavage activity by sterically hindering the formation of a cleavage-compatible state[13,14]. We observe that *Tt*Ago's activity and discrimination efficiency depend sensitively on the (i) position of the mismatched pair along the guide, (ii) buffer composition, (iii) guide concentration, (iv) guide length, (v) incubation temperature and time, and (vi) target sequence. *Tt*Ago appears to discriminate best between target and off-target in the presence of a mismatch at or around the cleavage site located between guide nucleotides 10 and 11. Optimally, the buffer should contain $[Mg^{2+}] \geq 8$ mM, 0.8 M betaine, and 1.4 mM dNTPs. The ssDNA guides should be 15-16nt in length with their concentration exceeding *Tt*Ago's concentration; and the incubation temperature should exceed the target dsDNA melting temperature. NAVIGATER is amenable to multiplexing and can concurrently enrich for multiple MAs while operating with different guides.

We demonstrate NAVIGATER's ability to enrich the fraction of cancer biomarkers such as *KRAS, BRAF,* and *EGFR* mutants in various samples. For example, NAVIGATER increased *KRAS* G12D fraction from 0.5% to 30% (60 fold) in a blood sample from a pancreatic cancer patient. The presence of 0.5% *KRAS* G12D could not be detected with Sanger sequencer or PNA-PCR. However after NAVIGATER pre-processing, both the Sanger sequencer and PNA-PCR readily identified the presence of *KRAS* G12D. Additionally, NAVIGATER combined with PNA-LAMP detects low fraction (0.1%) mutant RNA alleles and NAVIGATER combined with PNA-LAMP enables genotyping at the point of care and in resource-poor settings. NAVIGATER improves the detection limit of XNA-PCR by more than 10 fold, enabling detection of rare alleles with frequencies as low as 0.01%.

NAVIGATER differs from previously reported rare allele enrichment methods[8-10,34-38] in several important ways (**Table 3**). First, NAVIGATER is versatile. In contrast to CRISPR-Cas9[8-10] and restriction enzymes[34], *Tt*Ago does not require a PAM motif or a specific recognition site. A gDNA can be designed to direct *Tt*Ago to cleave any desired target. Second, *Tt*Ago is a multi-turnover enzyme[15]; a single *Tt*Ago-guide complex can cleave multiple targets. In contrast, CRISPR-Cas9 is a single turnover nuclease[39]. Third, whereas CRISPR-Cas9 exclusively cleaves DNA, *Tt*Ago cleaves both DNA and RNA targets with single nucleotide precision. Hence, NAVIGATER can enrich for both rare DNA alleles and their associated exosomal RNAs[40], further increasing assay sensitivity. Fourth, *Tt*Ago is robust, operates over a broad temperature range (66-86 ^{o}C) and unlike PCR-based enrichment methods, such as COLD-PCR[36] and blocker-PCR[37,38], does not require tight temperature control. Moreover, NAVIGATER can complement PCR-based enrichment methods. Fifth, *Tt*Ago is more specific than thermostable duplex-specific nuclease (DSN)[35]. Since DSN non-specifically cleaves all dsDNA, DSN-based assays require tight controls of probe concentration and temperature to avoid non-specific hybridization and cleavage of the rare nucleic acids of interest. Most importantly, as we have demonstrated, NAVIGATER is compatible with many downstream genotyping analysis methods such as ddPCR, PNA-PCR, XNA-PCR, and sequencing. Last but not least, NAVIGATER can operate with isothermal amplification methods such as LAMP, enabling integration of enrichment with genotyping for use in resource poor settings.

In the future, we will design panels of DNA guides to enable NAVIGATER in combination with downstream detection methods (including next-generation sequencing) to detect MAs indicative of various types of cancer. NAVIGATER could be leveraged to help detect other rare MA such as genetic disorders in fetal DNA and drug resistant bacteria.

Acknowledgements

Dr. Robert M. Greenberg helped with PAGE electrophoresis. Dr. Jennifer E. Phillips-Cremins provided us with access to gel imager. Dr. Changchun Liu provided helpful comments early in this project. Stephanie Yee and Taylor Black assisted with ddPCR. This work is supported by the NIH NCI 1R21CA227056-01 to the University of Pennsylvania, and by grants from the Netherlands Organization of Scientific Research (NWO-ECHO 711013002 and NWO-TOP 714015001) to J.v.d.O.

Author contributions

J.S. and H.H.B. conceived the project and designed the experiments. J.W.H. expressed and purified *Tt*Ago protein. J.S., J.W.H., J.P. and L.T.A. carried out the experiments. M.G.M. and J.C. assisted, respectively, with PNA-PCR and XNA-PCR. N.B. and J.E.T. extracted cfDNA from patients' blood and quantified *KRAS* G12 mutations with ddPCR. J.E.T. extracted RNA from cell lines. N.B. and J.E.T. assisted with ddPCR experiments. M.S. and J.M. cultured cell lines. E.C. supervised N.B., J.E.T., M.S., and J.M. and advised on patient samples and ddPCR experiments. J.S., J.W.H., M.G.M., J.v.d.O., and H.H.B. analyzed the data and wrote the manuscript. All authors read and commented on the manuscript.

Competing interests

University of Pennsylvania and Wageningen University have applied for a patent on NAVIGATER with J.S., J.W.H., M.G.M., J.v.d.O., and H.H.B. listed as co-inventors.

Human Subjects

This study was approved by Penn Institutional Review Board (IRB PROTOCOL #: 822028)

Methods

*Tt*Ago expression and purification

*Tt*Ago gene, codon-optimized for *E.coli* Bl21 (DE3), was inserted into a pET-His6 MBP TEV cloning vector (Addgene plasmid # 29656) using ligation-independent cloning. The *Tt*Ago protein was expressed in *E.coli* Bl21(DE3) Rosetta™ 2 (Novagen). Cultures were grown at 37 ^{o}C in Lysogeny broth medium containing 50 μg ml^{-1} kanamycin and 34 μg ml^{-1} chloramphenicol untill an OD_{600nm} of 0.7 was reached. *Tt*Ago-expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. During the expression, cells were incubated at 18 degrees for 16 hours with continuous shaking. Cells were harvested by centrifugation and lysed in buffer containing 20 mM Tris-HCl pH 7.5, 250 mM NaCl, 5mM

imidazole, supplemented with EDTA-free protease inhibitor cocktail tablet (Roche). The soluble fraction of the lysate was loaded on a nickel column (HisTrap Hp, GE healthcare). The column was extensively washed with buffer containing 20 mM Tris-HCl pH 7.5, 250 mM NaCl and 30 mM imidazole. Bound proteins were eluted by increasing the concentration of imidazole in the wash buffer to 250 mM. The eluted protein was dialysed at 4^{o}C overnight against 20 mM HEPES pH 7.5, 250 mM KCl, and 1 mM dithiothreitol (DTT) in the presence of 1 mg TEV protease (expressed and purified as previously described[41]) to cleave the His6-MBP tag. Next, the cleaved protein was diluted in 20 mM HEPES pH 7.5 to lower the final salt concentration to 125 mM KCl. The diluted protein was applied to a heparin column (HiTrap Heparin HP, GE Healthcare), washed with 20mM HEPES pH 7.5, 125 mM KCl and eluted with a linear gradient of 0.125-2 M KCl. Next, the eluted protein was loaded onto a size exclusion column (Superdex 200 16/600 column, GE Healthcare) and eluted with 20 mM HEPES pH7.5, 500 mM KCl and 1 mM DTT (**Supplementary Fig. 24**). Purified *Tt*Ago protein was diluted in a size exclusion buffer to a final concentration of 5 µM. Aliquots were flash frozen in liquid nitrogen and stored at -80^{o}C.

*Tt*Ago – based cleavage assays

5'-Phosphorylated DNA guides and Ultramer® ssDNA and ssRNA targets (100 nt) were synthesized by IDT (Coralville, IA). For ssDNA and ssRNA cleavage experiments, purified *Tt*Ago, DNA guides, and ssDNA or ssRNA targets were mixed with *Tt*Ago and guides at the ratios indicated in the buffers listed in **Table 1** and incubated at the indicated temperatures. Reactions were terminated by adding 1 µL proteinase K (Qiagen, Cat. No. 19131) solution, followed by 15 min incubation at 56°C. Samples were then mixed with 2X loading buffer (95% (de-ionized) formamide, 5mM EDTA, 0.025% SDS, 0.025% bromophenol blue and 0.025% xylene cyanol) and heated for 10 min at 95°C before the samples were resolved on 15% denaturing polyacrylamide gels (7M Urea). Gels were stained with SYBR gold Nucleic Acid Gel Stain (Invitrogen) and nucleic acids were visualized using a BioRad Gel Doc XR+ imaging system. For dsDNA cleavage, *TtAgo* and guides were pre-incubated in LAMP Buffer 3 (**Table 1**) at 75 °C for 20 min.

Cell-free DNA (cfDNA) and RNA Samples

Patient cfDNA samples. All six blood samples (**Table 2**) were obtained from patients with metastatic pancreatic cancer who had provided informed consent under the IRB-approved protocol (UPCC 02215, IRB# 822028). cfDNA was extracted with QIAamp® Circulating Nucleic Acid kit

(Qiagen, Valencia, CA, USA). Subsequently, the extracted cfDNA was qualified and quantified with multiplex ddPCR (Raindance).

RNA samples. Total RNA was extracted with RNeasy® mini kit (Qiagen, Valencia, CA, USA) per manufacturer's protocol from Human cancer cell lines U87-MG (WT *KRAS mRNA*) and ASPC1 (*KRAS* G12D mRNA) and quantified with ddPCR.

cfDNA pre-amplification was carried out in 50-µL reaction volumes using 20 ng of cfDNA, 1 × Q5 Hot Start High-Fidelity Master Mix (New England Biolabs, Ipswich, MA), and 100 nM each of forward and reverse *KRAS* 80 bp-PCR primers (**Supplementary Table 1**). Reaction mixes without DNA were included as no-template (negative) controls (NTCs). Nucleic acids were preamplified with a BioRad Thermal Cycler (BioRad, Model CFD3240) with a temperature profile of 98°C for 3 minutes, followed by 30 cycles of amplification (98°C for 10 seconds, 63°C for 3 minutes, and 72°C for 30 seconds), and a final 72 °C extension for 2 minutes.

mRNA pre-amplification was performed in 50-µL reactions using 30 ng of total RNA, 1 × Q5 Hot Start High-Fidelity Master Mix (New England Biolabs, Ipswich, MA), 100 nM each of forward and reverse *KRAS* 295 bp-PCR primers (**Supplementary Table 1**), and 1 µL reverse transcriptase (Invitrogen, Carlsbad, CA). The reaction mix was incubated at 55 °C for 30 minutes and 98°C for 3 minutes, followed by 30 cycles of amplification (93°C for 15 seconds, 62°C for 30 seconds, and 72°C for 30 seconds), and a final 72°C extension for 4 minutes.

Mutation enrichment (NAVIGATER)

The same setup as for synthetic dsDNA cleavage was used for cf-ctDNA and mutant mRNA enrichment. *TtAgo,* S-guide, and AS-guide were mixed in 1: 10: 10 ratio (1.25 µM *Tt*Ago, 12.5 µM S-guide, 12.5 µM AS-guide) in the Buffer 3 and pre-incubated at 75 °C for 20 min. Samples consisted of 2 µL preamplified PCR or RT-PCR products were added after pre-incubation of *Tt*Ago and guides. The reaction mixes were incubated at 83°C for 1 hour. The enriched products were diluted 10^4 fold before downstream mutation analysis or second-round enrichment. For second-round enrichment, the protocol outlined above was repeated.

NAVIGATER combined with downstream mutation detection methods

Droplet digital PCR (ddPCR). ddPCR was carried out with the RainDrop Digital PCR system (RainDance Technologies, Inc.) to verify mutation abundance before and after *Tt*Ago enrichment. 2

µL of the 10^4-fold diluted, *Tt*Ago-treated sample was added to each 30-µL dPCR. dPCRs contained 1× TaqMan Genotyping Master Mix (Life Technologies), 400 nM *KRAS* 80bp-PCR primers, 100 nM *KRAS* wild-type target probe, 100 nM *KRAS* mutant target probe (**Supplementary Table 1**), and 1× droplet stabilizer (RainDance Technologies, Inc.). Emulsions of each reaction were prepared on the RainDrop Source instrument (RainDance Technologies, Inc.) to produce 2 to 7 million, 5-pL-volume droplets per 25-µL reaction volume. Thereafter, the emulsions were placed in a thermal cycler to amplify the target and generate signal. The temperature profile for amplification consisted of an activation step at 95°C for 10 minutes, followed by 45 cycles of amplification [95°C for 15 seconds and 60°C for 45 seconds]. Reaction products were kept at 4°C before placing them on the RainDrop Sense instrument (RainDance Technologies, Inc.) for signal detection. RainDrop Analyst (RainDance Technologies, Inc.) was used to determine positive signals for each allele type. Gates were applied to regions of clustered droplets to define positive hits for each allele, according to the manufacturer's instructions.

PNA-PCR. PNA-PCR was performed in 20-µL reaction volumes, containing 4.5 µL of the 10^4-fold diluted *Tt*Ago-treated products, 1 × Q5 Hot Start High-Fidelity Master Mix (New England Biolabs, Ipswich, MA), 0.5 µL of EvaGreen fluorescent dye (Biotium, Hayward, CA), 500 nM *KRAS* PNA clamp (**Supplementary Table 1**), and 100 nM each of forward and reverse *KRAS* 80 bp-PCR primers. Reactions were amplified with a BioRad Thermal Cycler (BioRad, Model CFD3240) with a temperature profile of 98 °C for 3 minutes, followed by 40 cycles of amplification (98 °C for 10 seconds, 63 °C for 3 minutes, and 72 °C for 30 seconds).

Sanger sequencing. RNA extracted from cell lines were pre-amplified by *KRAS* 295 bp-PCR primers as described above and treated by *Tt*Ago mutation enrichment system. 2 µL of the 10^4-fold diluted, *Tt*Ago-treated sample was amplified by 295 bp PCR protocol (the same as 295 bp RT-PCR protocol without a reverse transcription step) for 30 cycles. PCR products were checked for quality and yield by running 5 µl in 2.2% agarose Lonza FlashGel DNA Cassette and processed for Sanger sequencing at Penn Genomic Analysis Core.

POC mutation detection. PNA-LAMP (SMAP-2) was prepared in 20-µL reaction volumes according to previously described protocol[18]. The reaction mix contained 2 µL of the 10^4-fold diluted *Tt*Ago-treated products (same as used for Sanger sequencing), 1 × LAMP buffer 3 (Eiken LAMP buffer), 1 µL *Bst* DNA polymerase (from Eiken DNA LAMP kit), 2.5 µL of BART reporter (Lot: 1434201; ERBA Molecular, UK)[20], *KRAS* PNA clamp and LAMP primers (sequences and concentrations

listed in **Supplementary Table 1**). The prepared reaction mixtures were injected into reaction chambers of our custom made multifunctional chip[42,43]. The inlet and outlet ports were then sealed with transparent tape (3M, Scotch brand cellophane tape, St. Paul, MN) and the chip was placed in our portable Smart-Connected Cup and processed according to previously described protocol[20].

Multiplexed NAVIGATER and XNA-PCR.

Single-plex and Multiplexed pre-amplification: Single-plex and Triplex PCR were carried out with mutation detection kit (DiaCarta, Inc[19]). The 10-µL reaction mixture contains 60 ng of cfDNA (reference standard that includes various MAs, Horizon Discovery, HD780), 1 × PCR Master Mix, 1 µL of either single or mixed PCR primers (1: 1: 1) for targets of interest. Nucleic acids were pre-amplified with a BioRad Thermal Cycler (BioRad, Model CFX96) with a temperature profile of 95 °C for 5 minutes, followed by 35 cycles of amplification (95 °C for 20 seconds, 70 °C for 40 seconds, 60 °C for 30 seconds, and 72 °C for 30 seconds), and a final 72 °C extension for 2 minutes.

Multiplexed enrichment: Guides (1: 1: 1) for targets of interest were mixed with *TtAgo* in 10: 1 ratio (12.5 µM S-guides, 12.5 µM AS-guides, 1.25 µM *Tt*Ago) in Buffer 3 and pre-incubated at 75 °C for 20 min. Samples consisted of 1 µL pre-amplified singleplex or triplex PCR products mixed with pre-incubated *Tt*Ago – guide complexes. The reaction mixes were incubated at 83°C for 1 hour. The products with and without *Tt*Ago treatment were resolved on 15% denaturing polyacrylamide gels (7M Urea). The products were diluted 10^7 fold before downstream mutation analysis.

XNA-PCR: NAVIGATER products were tested by mutation detection method XNA-PCR[19] (DiaCarta, Inc.). XNA-PCR was carried out for individual mutants in 10-µL reaction volumes, containing 3 µL of the 10^7-fold diluted NAVIGATER products, 1 × PCR Master Mix, 1 µL of PCR primer/probe mix, and 1 µL of XNA clamp. Reactions were amplified with a BioRad Thermal Cycler (BioRad, Model CFX96) with a temperature profile of 95 °C for 5 minutes, followed by 45 cycles of amplification (95 °C for 20 seconds, 70 °C for 40 seconds, 60 °C for 30 seconds, and 72 °C for 30 seconds).

REFERENCES

1. Barrangou, R. & Doudna, J.A. Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* **34**, 933-941 (2016).
2. Komor, A.C., Badran, A.H. & Liu, D.R. CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell* **168**, 20-36 (2017).
3. Wu, W.Y., Lebbink, J.H.G., Kanaar, R., Geijsen, N. & van der Oost, J. Genome editing by natural and engineered CRISPR-associated nucleases. *Nat Chem Biol* **14**, 642-651 (2018).
4. East-Seletsky, A.*, et al.* Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270-+ (2016).
5. Gootenberg, J.S.*, et al.* Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* **356**, 438-+ (2017).
6. Chen, J.S.*, et al.* CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* **360**, 436-+ (2018).
7. Gootenberg, J.S.*, et al.* Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **360**, 439-+ (2018).
8. Gu, W.*, et al.* Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol* **17**(2016).
9. Lee, S.H.*, et al.* CUT-PCR: CRISPR-mediated, ultrasensitive detection of target DNA using PCR. *Oncogene* **36**, 6823-6829 (2017).
10. Aalipour, A.*, et al.* Deactivated CRISPR Associated Protein 9 for Minor-Allele Enrichment in Cell-Free DNA. *Clin Chem* **64**, 307-316 (2018).
11. Hegge, J.W., Swarts, D.C. & van der Oost, J. Prokaryotic Argonaute proteins: novel genome-editing tools? *Nat Rev Microbiol* **16**, 5-11 (2018).
12. Swarts, D.C.*, et al.* The evolutionary journey of Argonaute proteins. *Nature Structural & Molecular Biology* **21**, 743-753 (2014).
13. Wang, Y.L.*, et al.* Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921-U972 (2008).
14. Wang, Y.L.*, et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754-U753 (2009).
15. Swarts, D.C.*, et al.* DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258-+ (2014).
16. Taly, V.*, et al.* Multiplex Picodroplet Digital PCR to Detect KRAS Mutations in Circulating DNA from the Plasma of Colorectal Cancer Patients. *Clin Chem* **59**, 1722-1731 (2013).
17. Choi, J.J.*, et al.* PNA-mediated Real-Time PCR Clamping for Detection of EGFR Mutations. *B Korean Chem Soc* **31**, 3525-3529 (2010).
18. Tatsumi, K.*, et al.* Rapid Screening Assay for KRAS Mutations by the Modified Smart Amplification Process. *Journal of Molecular Diagnostics* **10**, 520-526 (2008).
19. http://diacarta.com.
20. Song, J.Z.*, et al.* Smartphone-Based Mobile Detection Platform for Molecular Diagnostics and Spatiotemporal Disease Mapping. *Anal Chem* **90**, 4823-4831 (2018).
21. Sheng, G.*, et al.* Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *P Natl Acad Sci USA* **111**, 652-657 (2014).
22. Wang, Y.L., Sheng, G., Juranek, S., Tuschl, T. & Patel, D.J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209-U234 (2008).

23. Adamczak, B., Kogut, M. & Czub, J. Effect of osmolytes on the thermal stability of proteins: replica exchange simulations of Trp-cage in urea and betaine solutions. *Phys Chem Chem Phys* **20**, 11174-11182 (2018).
24. Henke, W., Herdel, K., Jung, K., Schnorr, D. & Loening, S.A. Betaine improves the PCR amplification of GC-rich DNA sequences. *Nucleic Acids Research* **25**, 3957-3958 (1997).
25. Salomon, W.E., Jolly, S.M., Moore, M.J., Zamore, P.D. & Serebrov, V. Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. *Cell* **162**, 84-95 (2015).
26. Joseph, T.T. & Osman, R. Convergent Transmission of RNAi Guide-Target Mismatch Information across Argonaute Internal Allosteric Network. *Plos Comput Biol* **8**(2012).
27. Swarts, D.C.*, et al.* Autonomous Generation and Loading of DNA Guides by Bacterial Argonaute. *Mol Cell* **65**, 985-+ (2017).
28. Phelps, C., Lee, W., Jose, D., von Hippel, P.H. & Marcus, A.H. Single-molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions. *P Natl Acad Sci USA* **110**, 17320-17325 (2013).
29. https://www.idtdna.com/calc/analyzer.
30. Bettegowda, C.*, et al.* Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. *Sci Transl Med* **6**(2014).
31. Oxnard, G.R.*, et al.* Noninvasive Detection of Response and Resistance in EGFR-Mutant Lung Cancer Using Quantitative Next-Generation Genotyping of Cell-Free Plasma DNA. *Clinical Cancer Research* **20**, 1698-1705 (2014).
32. Aggarwal, C.*, et al.* Clinical Implications of Plasma-Based Genotyping With the Delivery of Personalized Therapy in Metastatic Non-Small Cell Lung Cancer. *JAMA Oncol* (2018).
33. Tsiatis, A.C.*, et al.* Comparison of Sanger Sequencing, Pyrosequencing, and Melting Curve Analysis for the Detection of KRAS Mutations Diagnostic and Clinical Implications. *Journal of Molecular Diagnostics* **12**, 425-432 (2010).
34. Bielas, J.H. & Loeb, L.A. Quantification of random genomic mutations. *Nat Methods* **2**, 285-290 (2005).
35. Song, C.*, et al.* Elimination of unaltered DNA in mixed clinical samples via nuclease-assisted minor-allele enrichment. *Nucleic Acids Research* **44**(2016).
36. Li, J.*, et al.* Replacing PCR with COLD-PCR enriches variant DNA sequences and redefines the sensitivity of genetic testing. *Nat Med* **14**, 579-584 (2008).
37. Kim, H.S.*, et al.* Predictive efficacy of low burden EGFR mutation detected by next-generation sequencing on response to EGFR tyrosine kinase inhibitors in non-small-cell lung carcinoma. *PLoS One* **8**, e81975 (2013).
38. Wu, L.R., Chen, S.X., Wu, Y., Patel, A.A. & Zhang, D.Y. Multiplexed enrichment of rare DNA variants via sequence-selective and temperature-robust amplification. *Nat Biomed Eng* **1**, 714-723 (2017).
39. Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C. & Doudna, J.A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62-+ (2014).
40. Krug, A.K.*, et al.* Improved EGFR mutation detection using combined exosomal RNA and circulating tumor DNA in NSCLC patient plasma. *Ann Oncol* **29**, 700-706 (2018).
41. Tropea, J.E., Cherry, S. & Waugh, D.S. Expression and Purification of Soluble His6-Tagged TEV Protease. in *High Throughput Protein Expression and Purification: Methods and Protocols* (ed. Doyle, S.A.) 297-307 (Humana Press, Totowa, NJ, 2009).
42. Song, J.Z.*, et al.* Molecular Detection of Schistosome Infections with a Disposable Microfluidic Cassette. *Plos Neglect Trop D* **9**(2015).

43. Song, J.Z., *et al.* Instrument-Free Point-of-Care Molecular Detection of Zika Virus. *Anal Chem* **88**, 7289-7294 (2016).

Tables and Figures

Table 1. Buffer compositions

Buffer 1 ThermoPol® Reaction Buffer (1X, NEB)	Buffer 2 Isothermal Amplification Buffer (1X, NEB)	Buffer 3 Eiken buffer (1X)	Buffer S[13] Buffer of Swarts et.al.
20 mM Tris-HCl 10 mM $(NH_4)_2SO_4$ 10 mM KCl 2 mM $MgSO_4$ 0.1% Triton® X-100 (pH 8.8 @ 25°C)	20 mM Tris-HCl 10 mM $(NH_4)_2SO_4$ 50 mM KCl 2 mM $MgSO_4$ 0.1% Tween® 20 (pH 8.8 @ 25°C)	20 mM Tris-HCl 10 mM $(NH4)_2SO_4$ 10 mM KCl 8 mM $MgSO_4$ 0.1% Tween20 0.8 M Betaine 1.4 mM dNTPs (pH 8.8 @ 25°C)	10 mM Tris-HCl 125 mM NaCl 2 mM $MgCl_2$ (pH 8.0 @ 25°C)

Table 2. The genotype and mutation frequency of 6 samples from pancreatic cancer patients*

Patient number	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**
Genotype and mutation fraction**	ND**	ND**	G12R 3%	G12V 5%	G12D 0.5%	G12D 20%

* Samples were analyzed with standard ddPCR protocol[16].
** Not detected, possibly due to mutant allele not present.
All samples contain similar numbers of WT-*KRAS* (s.d. <10%). See **Supplemental Fig. 22b**.

Table 3. Comparison of rare allele enrichment methods

	dCas9-based method[10]	DASH[8] and Cut-PCR[9]	NAVIGATER	NaME-PrO[35]	COLD-PCR[36]	Blocker (PNA or DNA)-PCR[37,38]
Enzyme	deactivated CRISPR Cas9 (high cost)	CRISPR Cas proteins (low cost)	*Tt*Ago (low cost)	DSN (high cost)	N/A*	N/A
Guide/blocker	~120 nt sgRNA (high cost)	~120 nt sgRNA (high cost)	15/16 nt DNA (low cost)	20-25 nt DNA (low cost)	N/A	PNA (high cost) DNA (low cost)
PAM site requirement	Yes	Yes	No	No	No	No
Multi-turnover enzyme	No	No	Yes	Yes	N/A	N/A
Target	DNA	DNA	DNA & RNA	DNA	DNA	DNA
Incubation time	~7 h	2 h	<1 h	22 min	~2 h	~2 h
Tight temperature control	No (37°C)	No (37 °C)	No (66-86 °C)	Yes (98 °C, 67 °C)	Yes (thermal cycling)	Yes (thermal cycling)
Fraction detection limit	0.1% with allele-specific qPCR	0.01% with targeted deep sequencing	0.01% with XNA-PCR	Routinely 0.01% with HRM/Sanger sequencing	0.1%-0.5% with MALDI-TOF	<0.1%
Enrichment level	5-10 fold	5-10 fold	5-10 fold	5-20 fold	5-12 fold	N/A
Specificity	Medium	High	High	Medium	Medium	Medium
Multiplexing capability	Yes	Yes	Yes	Yes	Yes	Yes

* N/A – not applicable

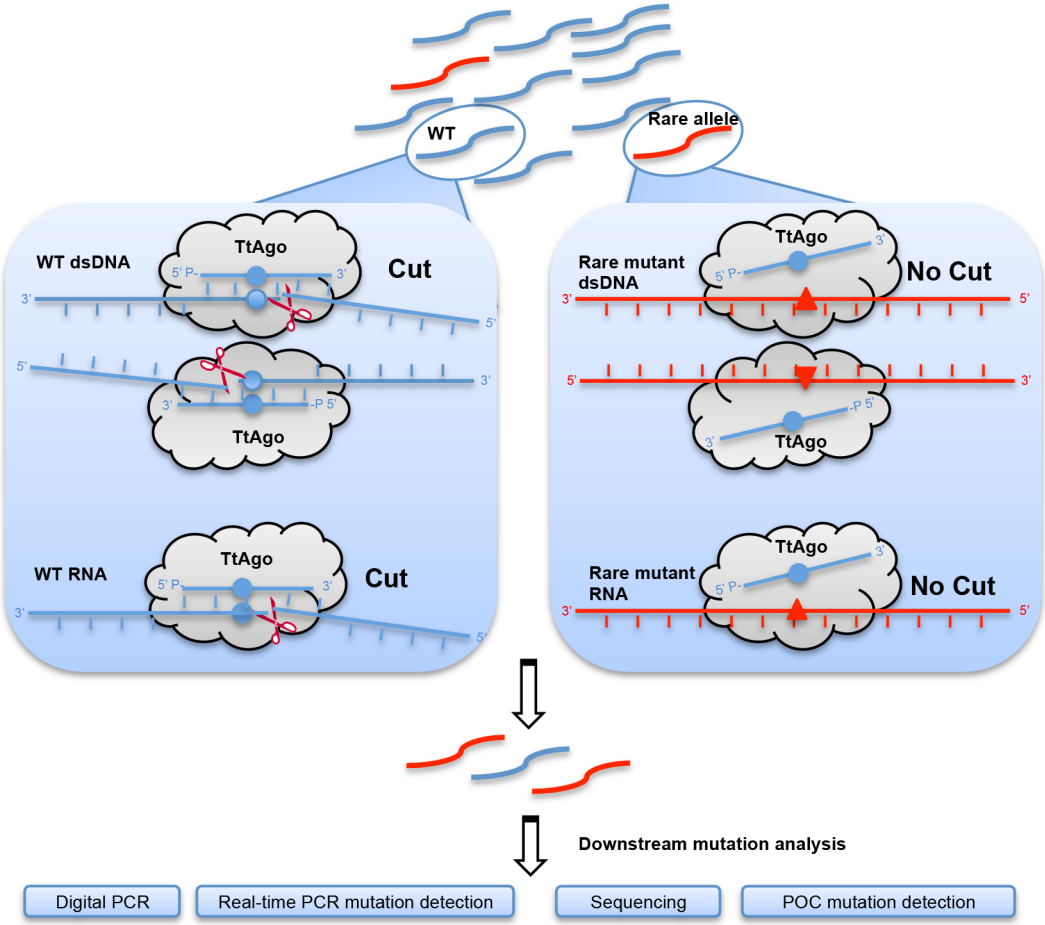

Figure 1: Argonaute-Based Enrichment Assay for Rare Alleles (NAVIGATER) cleaves guide-complementary targets at the cleavage site between guide positions g10 and g11 (counted from the 5' end) while sparing rare mutant alleles.

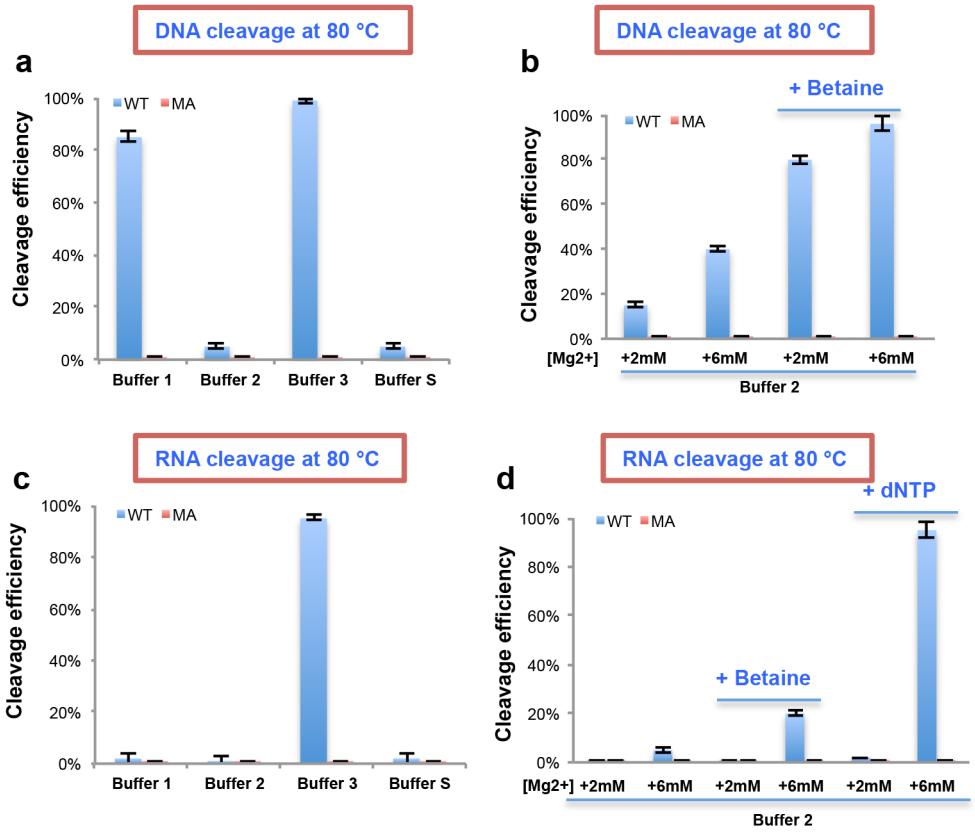

Figure 2: Cleavage efficiency (Γ) of WT *KRAS* and *KRAS* G12D DNA and RNA in Buffers 1, 2, 3, and S (Table 1) at 80°C. (a) DNA cleavage as a function of buffer composition. (b). DNA cleavage in Buffer 2 as a function of added [Mg^{2+}] in the absence and presence of betaine. (c) RNA cleavage as a function of buffer composition. (d) RNA cleavage in Buffer 2 as a function of added [Mg^{2+}] in the absence and presence of betaine and dNTP. All experiments were carried out with *KRAS* Sense (S) strand and 16 nt *KRAS*-S guide with a pair mismatch at position 12 (MP12). Incubation time 20 min. *Tt*Ago: guide: target =1: 0.2: 0.2. N=3.

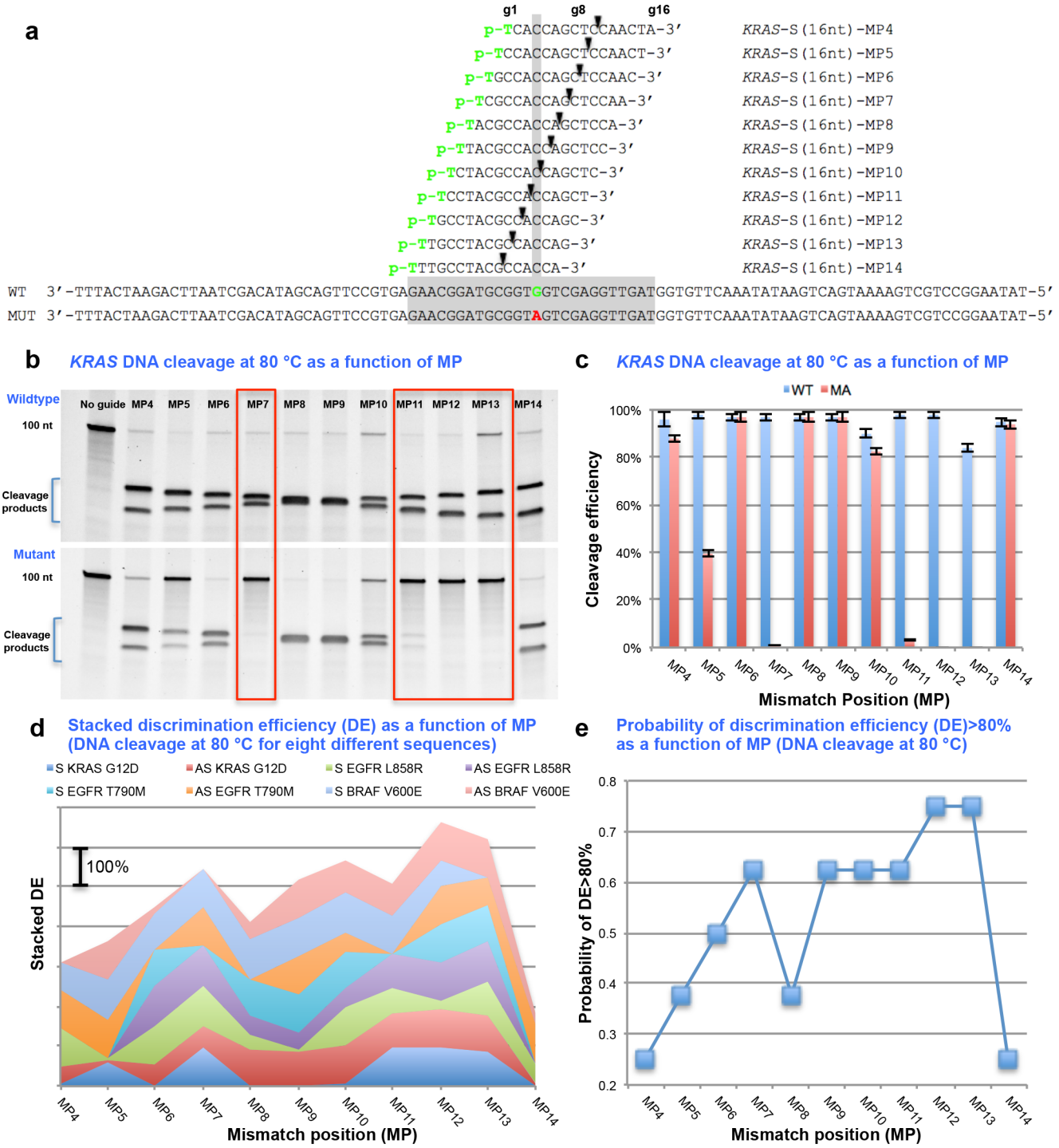

Figure 3: NAVIGATER's discrimination efficiency depends sensitively on guide-off target mismatch pair's position (MP). (**a**) Overview of the *KRAS* – S guide and S target sequences. The various guides vary in the position of the pair mismatch between S gDNA and S *KRAS* G12D. (**b**) Electropherograms (polyacrylamide urea gel) of NAVIGATER (80 °C, 20 min) products of S WT *KRAS* DNA and S *KRAS* G12D DNA strands as functions of MP. (**c**) Cleavage efficiencies of S WT *KRAS* DNA and S *KRAS* G12D DNA as a function of MP. (**d**) Stacked discrimination efficiency (DE) (difference between cleaving efficiency of WA and MA) of S and AS *KRAS* G12D, *EGFR* L858R, *EGFR* T790M, and *BRAF* V600E DNA as a function of MP. (**e**) Probability of DE >80% as a function of MP for all cases examined in (d). All experiments were carried out with short guides (15/16 nt) in Buffer 3 at 80 °C. *Tt*Ago: guide: target =1: 0.2: 0.2. N=3. The mismatch types and DE values are summarized in **Supplementary Fig. 14**.

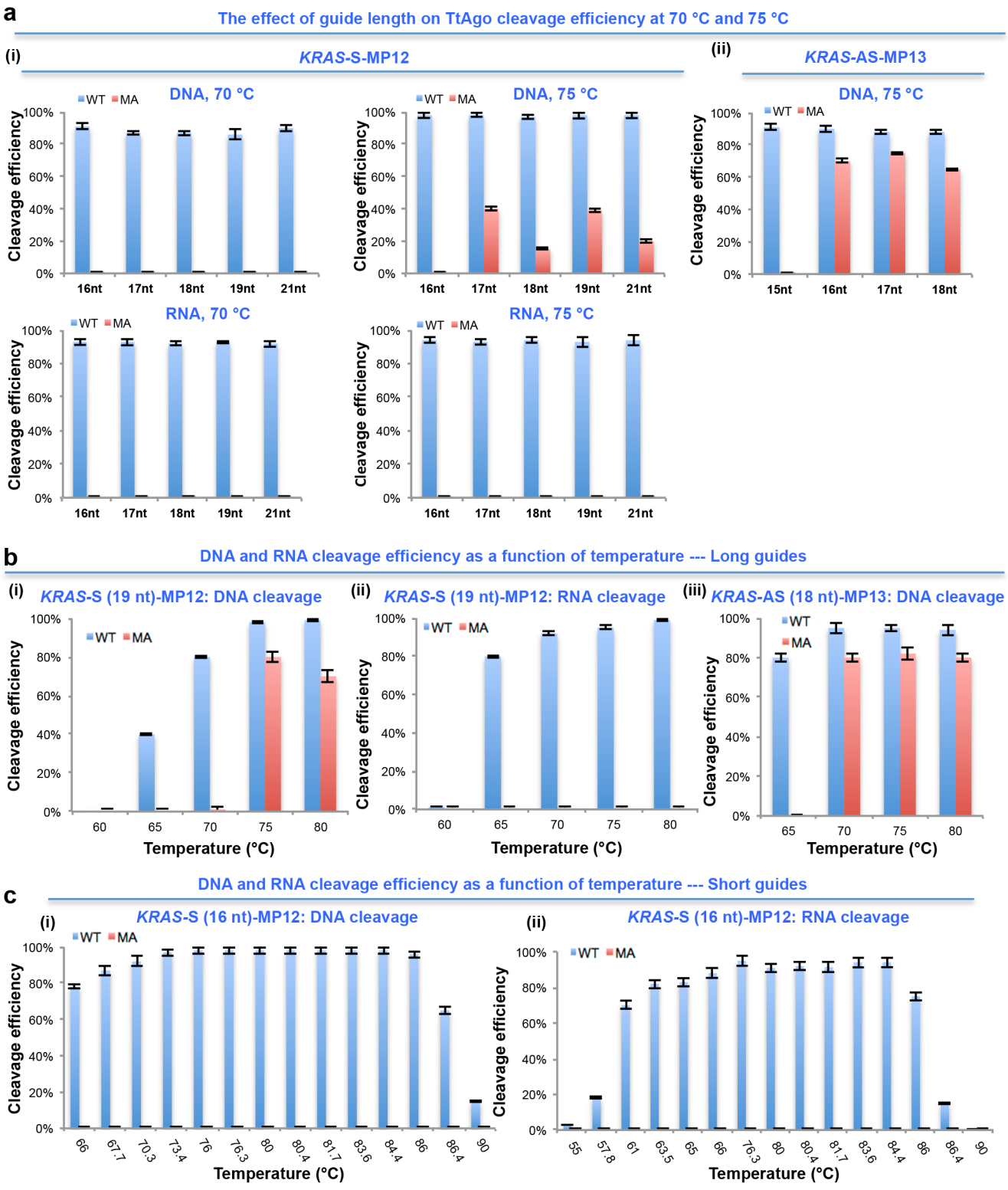

Figure 4: Short DNA guides (15/16 nt) provide a better discrimination between WT and MA. (**a**) Cleavage efficiency as a function of guide length at 70 °C and 75°C: (i) WT *KRAS* and *KRAS* G12D, S-DNA and RNA and (ii) AS WT *KRAS* and *KRAS* G12D. (**b**) Cleavage efficiency as a function of temperature with 18/19 nt long guide (i) WT *KRAS* and *KRAS* G12D S-DNA, (ii) WT *KRAS* and *KRAS* G12D RNA, and (iii) WT *KRAS* and *KRAS* G12D AS-DNA. (**c**) Cleavage efficiency of WT *KRAS* and *KRAS* G12D (i) S-DNA and (ii) RNA as functions of temperature (16 nt long guide). Buffer 3. *Tt*Ago: guide: target =1: 0.2: 0.2. N=3.

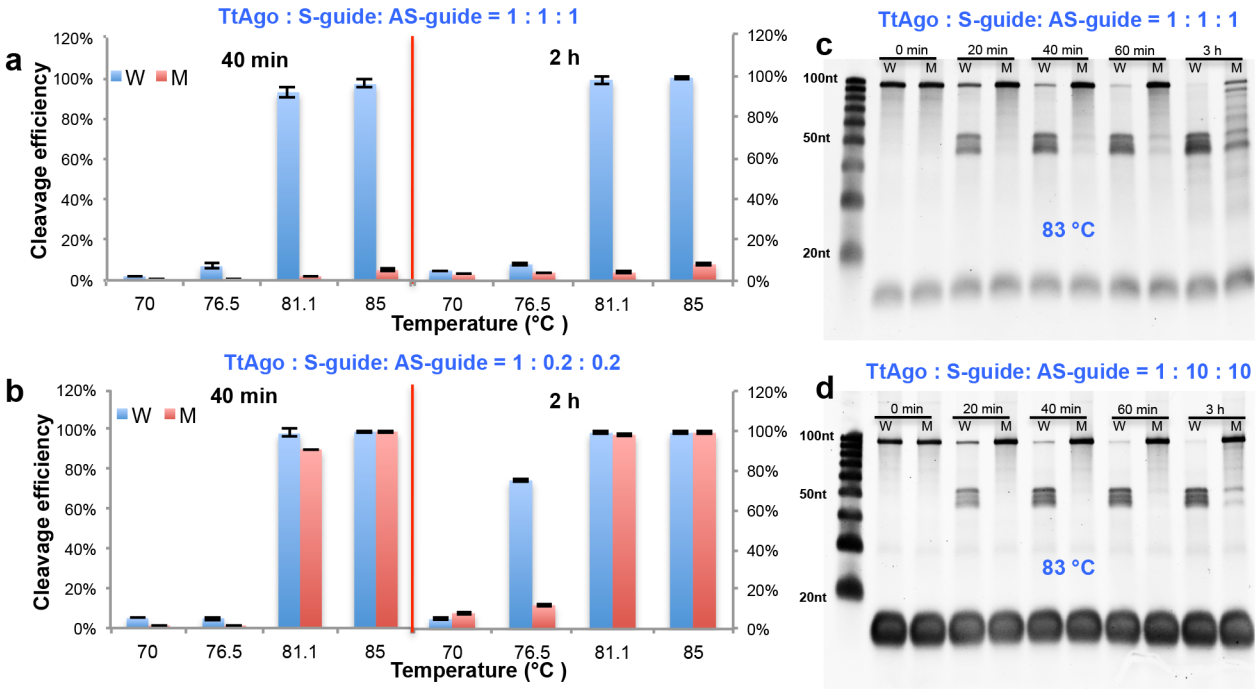

Figure 5: Excess guide concentration provides high discrimination efficiency. dsDNA *KRAS* and *KRAS* G12D cleavage efficiencies as functions of temperature: (**a**) *Tt*Ago/(S-guide)/(AS-guide) concentration ratio: 1: 1 : 1, 40 min and 2h incubation time; (**b**) *Tt*Ago/(S-guide)/(AS-guide) concentration ratio: 1: 0.2 : 0.2, 40 min and 2h incubation time. Electropherograms of NAVIGATER's products as a function of incubation (83 ^{o}C) time: (**c**) *Tt*Ago/(S-guide)/(AS-guide) ratio 1: 1: 1 and (**d**) *Tt*Ago/(S-guide)/(AS-guide) ratio 1: 10: 10. All experiments were carried out in Buffer 3 with *KRAS*-S (16nt)-MP12 and *KRAS*-AS (15nt)-MP13 guides. N=3.

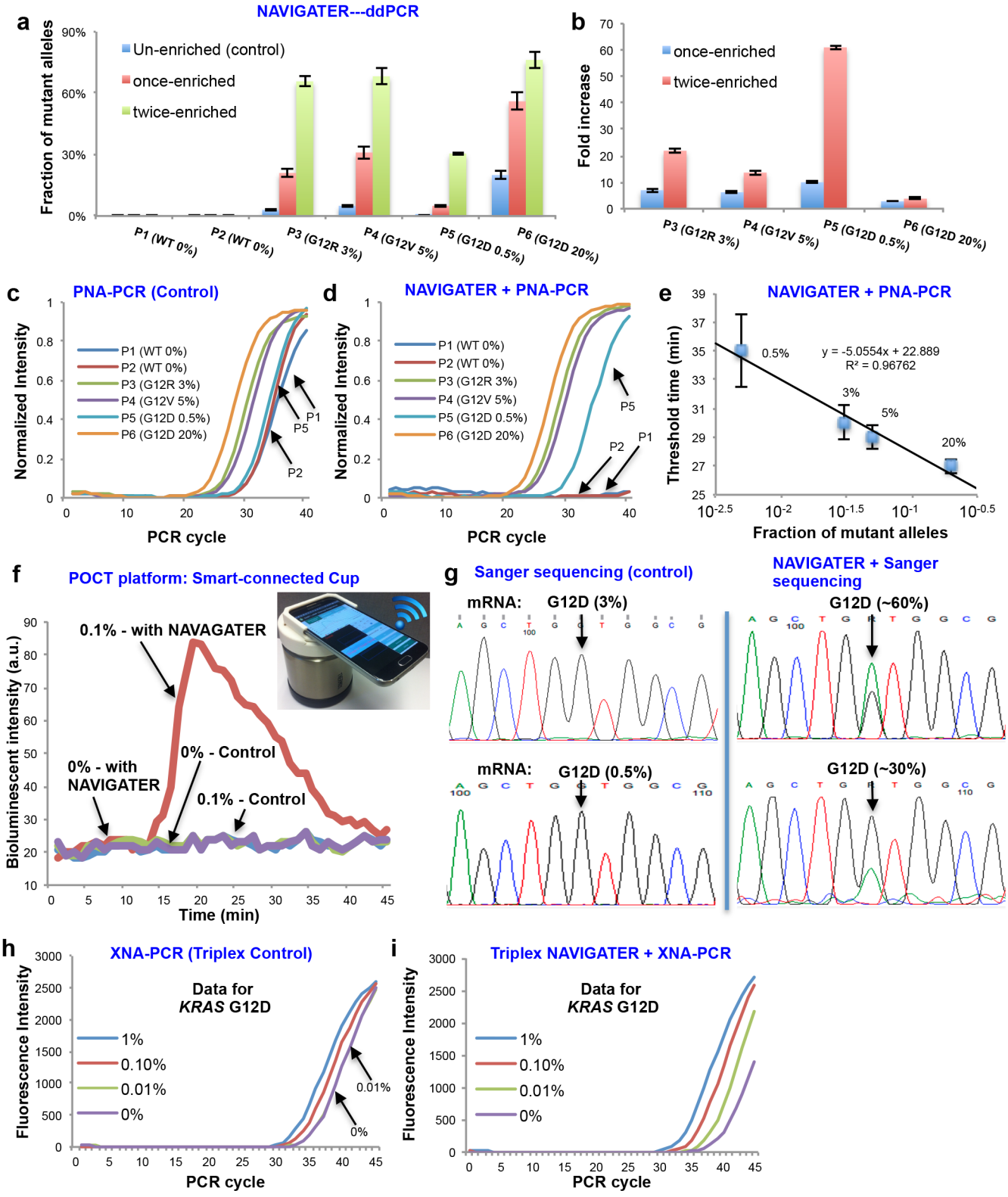

Figure 6: NAVIGATER enhances sensitivity of downstream detection methods. (a, b) ddPCR of samples from pancreatic cancer patients containing *KRAS* mutants (**Table 2**): (**a**) Fraction of droplets reporting mutant alleles. (**b**) Increase in mutant allele fraction after NAVIGATER enrichment. (**c, d, e**) PNA-PCR's amplification curves of pancreatic cancer patients' samples before (**c**) and after (**d**) NAVIGATER, and amplification threshold time as a function of mutant fraction (**e**). (**f**) PNA-LAMP of simulated RNA samples before and after NAVIGATER carried out with a minimally-instrumented, electricity-free Smart-Connected Cup (SCC)[20] (inset). (**g**) Sanger sequencing before and after

NAVIGATER when detecting simulated RNA samples. (**h, i**) XNA-PCR of samples containing various mutant concentrations without (h) and with (i) multiplexed NAVIGATER pre-treatment. All the controls were pre-processed with NAVIGATER in the absence of *Tt*Ago.