A-Star, an Argonaute-directed System for Rare SNV Enrichment and Detection

Qian Liu[1], Xiang Guo[1], Guanhua Xun[1], Zhonglei Li[1], Yuesheng Chong[1], Litao Yang[1], Hongxia Wang[2], Fengchun Zhang[3], Shukun Luo[1], Zixin Deng[1], Kai Li[4], Yan Feng[1,]*

Dr. Q. Liu, X. Guo, G. Xun, Z. Li, Y. Chong, Prof. L. Yang, Prof. S. Luo, Prof. Z. Deng, Prof. Yan Feng

State Key Laboratory of Microbial Metabolism, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

E-mail: yfeng2009@sjtu.edu.cn

Prof. H. Wang

Department of Oncology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200080, China

Prof. F. Zhang

Department of Oncology, Renji Hospital & Suzhou Kowloon Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200127 & Suzhou 215021, China

Dr. K. Li

GeneTalks Biotechnology Inc., Changsha, Hunan 410013, China

Key words:

single nucleotide variations (SNVs), A-Star (Ago-directed specific target enrichment), *Pyrococcus furiosus* Argonaute (*Pf*Ago), multiplex detection

Abstract: The detection of rare polymorphic alleles, especially single nucleotide variations (SNVs), is becoming increasingly relevant for the early diagnosis and monitoring of a variety of tumors. However, the broad clinical applicability of SNV detection has been limited by the low sensitivity, flexible target selection and

multiplexing offered by existing SNV detection methods. Here, we developed a simple but efficient single-tube PCR system, referred to as A-Star (Ago-directed specific target enrichment), that takes advantage of the thermophilic nature of *Pyrococcus furiosus* Argonaute (*Pf*Ago) to specifically cleave wild-type sequences during the DNA denaturation step, leading to progressive and rapid (~ 3 h) enrichment of scarce SNV-containing alleles. This approach enables the detection of rare SNVs at an attomolar sensitivity and can amplify allele fractions as low as 0.01% with an over 6000-fold efficiency. We further validated the A-Star system by multiplex detection of three rare oncogenic genes in complex genetic backgrounds. Finally, we successfully applied A-Star to detect oncogenic mutations in cancer patient tissue and blood samples, indicating that our method holds great promise for the diagnosis of various types of cancer as well as the genotyping of rare alleles for basic research.

The ability to detect genetic polymorphisms has revolutionized the molecular diagnosis and monitoring of tumors and may also shed light on the mechanisms underlying tumor onset and development[1]. The clinical detection methods for SNVs that are currently available involve either sequencing or PCR[2]. Compared with authentication by sequencing DNA fragments, allele-specific diagnostic PCR is not only simpler and timesaving but also practical and effective[3]. Therefore, the detection of extremely rare variant alleles within a complex mixture of DNA molecules has attracted increasing attention aimed at solving the technical problems associated with the strict requirements for both a precise single-nucleotide resolution and simple multiplex detection, especially in the detection of circulating tumor DNA in the plasma of early-stage patients[4]. Recently, a variety of SNV detection methods based on endonucleases with sequence-specific catalytic capabilities[5], particularly the enzymes of the CRISPR/Cas systems, have been developed[6]. These strategies are primarily based on protospacer-adjacent motif (PAM)-dependent Cas9 cleavage activity[7] or the protospacer flanking site (PFS)-dependent collateral cleavage activity of Cas12 and Cas13a[8]. Interestingly, when the Cas12/13-enzymes are coupled with recombinase polymerase amplification (RPA) in a single-tube assay, attomolar sensitivity can be achieved[9]. However, the major limitation of these approaches is the strict requirement for a PAM for Cas9 and a PFS for Cas13a in the vicinity of target sequence cleavage[9a, 10], which restricts the efficiency of these methods for the detection of a wide spectrum of oncogenes and requires tedious enzyme screening to target multiple genomic loci simultaneously[11].

One potential means of circumventing these limitations is to take advantage of the unique properties of Argonaute (Ago) proteins, which do not require a prerecognition motif and instead specifically cleave target nucleic acids according to the base pairing of guide-target strands[12]. A deep understanding and diligent exploration of the unique

properties through which guide nucleic acid-directed Ago recognizes target DNA (tDNA) may facilitate the development of new methods that enable the detection of rare SNVs with a high resolution and great precision[12c, 13]. Indeed, the *Tt*Ago-based method was developed for PAM-independent SNV enrichment, and can be easily combined with multiple downstream detection methods[14]. However, due to the incompatibility of the working temperatures of *Tt*Ago (70-80 °C) and PCR, repeated cleavage and PCR steps are needed for optimal enrichment efficiency. Here, we report the establishment of A-Star (Ago-directed specific target enrichment), a simple but efficient single-tube PCR system that takes advantage of hyperthermophilic Ago from *Pyroccocus furiosus* (*Pf*Ago)[12d, 15] to specifically cleave wild-type (WT) sequences during the DNA denaturation step, leading to progressive and rapid enrichment of rare SNV-containing alleles (**Figure 1**a). The guide DNAs (gDNAs) are designed and optimized to extend the precise complementation of WT single-stranded DNA (ssDNA) but not SNV alleles and to direct the *Pf*Ago-specific cleavage of unwinding double-stranded DNA (dsDNA) at 94 °C, thus enabling the elimination of a large fraction of the WT background. The subsequent annealing and strand-extension steps initiated by routine PCR primers and DNA polymerase favor the amplification of the spared SNV allele template and not the cleaved WT fragments. The catalytic elimination of WT DNA accompanied by the exponential amplification of the SNV alleles in each cycle of PCR led to enrichment of the SNVs in a manner that was both efficient and easy to perform. Coupling the A-Star approach with TaqMan real-time quantitative PCR (TaqMan PCR) or Sanger sequencing allows the detection of the enriched rare SNV sequence (full details are presented in the Methods).

As an initial presentation of the feasibility of using thermophilic Ago for PCR, we examined the temperature dependence of *Pf*Ago cleavage activity against ssDNA

targets. Previous studies indicated that *Pf*Ago prefers to bind a 16 nt single-stranded gDNA with 5'-phosphate and that the complex pairs with and cleaves the ssDNA target at matching gDNA positions between the 10th and 11th positions[12c, 13]. When the *Pf*Ago-gDNA complex was incubated with a synthesized *KRAS* ssDNA fragment of 60 nt at the indicated temperatures for 15 minutes, which is longer than the total cumulative time of the DNA denaturing step in PCR, we observed high DNA cleavage activities at temperatures ranging from 80 to 97°C but not below 70 °C (Figure S1, Supporting Information), indicating that *Pf*Ago catalysis is compatible with the DNA denaturing step of PCR at 94°C but is also favorable for continuous WT DNA cleavage at high temperature during the PCR thermal cycling. To avoid nonspecific extension initiated by the gDNA in subsequent PCR amplification for both the WT and SNV alleles, we modified each gDNA by adding an extra 3'-phosphate and verified its effectiveness by pre-PCR (Figure S2, Supporting Information). Together, the results of the following experiments indicated that *Pf*Ago cleavage and PCR amplification are compatible in a single reaction directed by a 16 nt gDNA in which both the 5'- and 3'-termini are modified by phosphate groups.

To address the precise cleavage of tDNA, the crucial concern in A-Star involves the design and selection of the gDNAs for the discrimination of SNVs. To achieve the goal of specifically cleaving wild-type (WT) ssDNA (unwinding double-stranded DNA (dsDNA) during the DNA denaturation step), we optimized the gDNA design using 60 nt of the synthesized *KRAS* ssDNA fragment of the WT or G12D allele as a representative cleavage target. Considering the potential effect of mismatch position and nucleotide type on DNA cleavage efficiency[12c, 13], we applied a systematic design for selecting the best gDNA hit for the highly specific cleavage of the WT sequence. We found that both WT *KRAS* and mutant *KRAS* containing an SNV encoding a G12D

substitution were cleaved indifferently when 16 nt gDNA sequences perfectly matching the WT sequence were incubated with *Pf*Ago and *KRAS* ssDNA at 94°C for 15 min (Figure S3, Supporting Information). This result suggests that the single-nucleotide difference in the gDNA for the WT and SNV alleles is not sufficient for discriminating DNA cleavage. To improve the discrimination between the WT DNA and sequences containing SNVs, we introduced an additional mismatched nucleotide of the gDNA at positions around its SNV pairing site and then evaluated the effect on cleavage efficiency. When a mismatch was introduced at position -4, -1, or +1 of the gDNA, corresponding to gDNA gM7, gM10 and gM11, respectively, for the SNV allele, *Pf*Ago preferred to catalyze the WT *KRAS* ssDNA, resulting in the lowest degree of SNV allele cleavage without compromising WT allele cleavage (Figure 1b and Figure S4, Supporting Information). For simplicity, we focused on double, continuous mismatches at positions -1 and +1 of the SNA site for further gDNA optimization in terms of nucleotide types. We observed that the mismatched nucleotide identity of the gDNA affects the discrimination performance of *Pf*Ago, but with a high dependency on the target sequence, not the specific nucleotides (G/C or A/T) (Figure S5a-d and S6a-d, Supporting Information). Collectively, these results indicated that gDNAs with double mismatches at positions around their SNV sites can direct the preferential cleavage of WT sequences, thereby allowing SNV selection. Considering the complexity of the gene sequence, we generally expand the gDNA candidate spectrum for the best hit from two gDNA groups with mismatched nucleotides at positions (-1, 0) and (0, +1) for the SNV allele in the following work to discriminate WT and rare SNV alleles.

After optimizing the gDNA candidates for each single strand of the WT DNA, we further determined the recombination effect of the paired gDNAs on the discriminating cleavage of a dsDNA target to completely eliminate the WT sequence in the PCR

system. When we used a pair of best hit gDNAs, the SNV variants of *KRAS* G12D and *PIK3CA* E545K remained uncleaved, whereas their corresponding WT counterparts were almost completely cleaved by incubation with *Pf*Ago at 94 °C for 15 min (Figure S5e-f and S6e-f, Supporting Information). When a pair of semibest hit gDNAs in which one best gDNA was mixed with a poorer hit gDNA was used, we did not observe clear discriminating cleavage of the dsDNA target. This result verified that one pair of the best hit gDNAs could direct *Pf*Ago to carry out discriminating cleavage of each individual strand of the unwinding dsDNA. Additionally, considering that the indel mutation is quite popular in oncogenes, we asked whether a pair of gDNAs matched to the WT sequence could discriminate between the WT and mutant. Taking *EGFR* and the its 15-nt deletion mutant as an experimental case, it was confirmed that *Pf*Ago can efficiently cleave WT dsDNA with a minimal effect on the mutant at 95 °C (Figure S7, Supporting Information). Taken together, these results demonstrated that *Pf*Ago is capable of effectively discriminating mutated dsDNA from WT dsDNA when directed by a pair of gDNAs that each contain two consecutive nucleotides mismatched with the SNV templates at position +1 or -1.

We next asked whether gDNA-directed *Pf*Ago targeting could be combined with a standard PCR amplification program for SNV enrichment and detection in a single reaction tube. We tested the effect of *Pf*Ago on the efficient cleavage of only the WT and not the SNV sequence by using the *KRAS* G12D WT and SNV alleles, respectively, in the presence of a pair of best hits gDNAs (Figure 1c). PCR was performed with amplification cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 20 s after preincubation, as suggested by the commercial protocol. Gel electrophoresis provides direct proof of discriminating cleavage since the products from WT are almost completely digested into two small fragments, whereas SNV alleles are enriched as

single, full-length DNA fragments (Figure 1c**)**. When a mixture of WT and SNV alleles at a ratio of 1:1 was used as a template, three bands appeared following the mixed reaction including the above two individual reactions. Sanger sequencing also provided support for the enrichment of SNV alleles, since the guanidine nucleotide peak from WT became a major peak, differing from the results for the unamplified sample, with two distinct peaks (Figure S8, Supporting Information). We further evaluated SNV allele enrichment quantitatively by TaqMan PCR. The results showed that more than ~80% of the amplicons came from the SNV templates (Figure 1d). Collectively, the results indicated that the PCR approach combined with the use of gDNA pair-directed *Pf*Ago led to SNV enrichment by eliminating WT template DNA in a single tube.

Seeking to increase the detection sensitivity for rare SNV alleles from a 1% variant allele fraction (VAF) to a 0.01% VAF, we systematically titrated the *Pf*Ago amount, the molar ratio between *Pf*Ago and gDNA, and the PCR thermal cycling conditions for a sample with a 1% VAF. We found that the *Pf*Ago concentration greatly affected SNV enrichment (Figure S9, Supporting Information**)**. As the *Pf*Ago concentration was increased to 30 nM with a ratio of 3:1 between *Pf*Ago and tDNA, the majority of WT alleles could be removed, leading to efficient rare SNV enrichment; when the concentration was over 30 nM and the molar ratio reached 4:1 or 5:1, the SNV enrichment efficiency decreased, which might have been caused by the strong binding ability of *Pf*Ago toward unspecific DNA and reduced the efficiency of tDNA concentration for catalytic cleavage. We also found that an increased ratio of 20:1 between gDNA and *Pf*Ago could enhance SNV enrichment (Figure S10, Supporting Information**)**, implying that gDNA with a high concentration could competitively bind *Pf*Ago and release the bound tDNA for further cleavage. Moreover, during the PCR procedure, 25~30 thermal cycles could efficiently enrich the rare alleles. Taken

together, the results showed that the optimal conditions were 30 nM *Pf*Ago with a *Pf*Ago:gDNA ratio of 1:20 and a 10 nM DNA sample. Moreover, optimum discrimination was obtained with a 25-cycle PCR program (Figure S11, Supporting Information).

Furthermore, to assess the possibility of continuous cleavage by *Pf*Ago in the PCR-coupling reaction for the efficient enrichment of rare SNV alleles, we compared A-Star consisting of single-tube *Pf*Ago-coupled PCR with uncoupled *Pf*Ago PCR involving a two-step reaction comprising 15 min of incubation of 1% *KRAS* G12D SNV alleles with the *Pf*Ago-gDNA complex at 94°C to cleave the WT template and then 25 cycles of PCR (**Figure 2**a). We observed a marked difference in the SNV enrichment efficiency between our single-tube *Pf*Ago-PCR approach and the uncoupled *Pf*Ago catalysis and PCR approach. The A-Star strategy increased the percentage of *KRAS* G12D SNV sequences in the final PCR products from 1% to 83% in comparison with an increase from 1% to only 6% in the uncoupled *Pf*Ago-PCR strategy, as assessed by TaqMan PCR (Figure 2b and Figure S12, Supporting Information) and Sanger sequencing (Figure S13, Supporting Information). The results suggest that continuous gDNA-directed *Pf*Ago cleavage of WT DNA during PCR cycles could enhance SNV allele enrichment significantly compared to uncoupled enzyme catalytic cleavage and PCR.

Next, we evaluated the limit of detection (LOD) of A-Star according to the differential WT/mutant ratio. *KRAS* G12D VAFs of 0.01%, 0.1%, and 1% were used for the analysis. Although the initial VAFs varied by 10-100-fold, the SNV allele enrichment ratio showed little difference, ranging from 60 to 70%, as detected by TaqMan PCR (Figure 2c). Remarkably, even for a sample with a VAF as low as 0.01%,

the SNV products could be enriched over 6000-fold (Figure 2d) and were readily detected by Sanger sequencing after the A-Star procedure (Figure S14, Supporting Information). Similar results were obtained at a VAF of 0.01% for samples containing a *PIK3CA* SNV or an *EGFR* indel mutation (Figure S15, Supporting Information).

To explore the clinical utility of A-Star, we further evaluated the LOD of A-Star using a set of commercial cell-free DNA (cfDNA) standards containing genomic DNAs with three available VAFs of 0.1%, 1%, and 5% for *KRAS* G12D, *PIK3CA* E545K and the *EGFR* indel. Since the low amount of cfDNA in mock samples was not directly detectable using the A-Star method (data not shown here), we added 30 cycles of pre-PCR amplification to increase the amount of cfDNA for the WT and SNV alleles. The inclusion of this step enabled the enrichment and detection of SNVs from samples containing only 33 ng of cfDNA with a VAF of as low as 0.1% for the SNVs (Figure S16, Supporting Information). In addition, we evaluated the mean cycle threshold (Ct) values of the samples with different VAFs and found an obvious linear relationship (Figure 2e), indicating a high potential for quantitative analysis of the VAFs of oncogenes in the early stage diagnosis or monitoring of oncogenes. All of these results for the mock cfDNA samples emphasize that A-Star offers impressive specificity for analyzing high-background DNA. Furthermore, we tested A-Star on DNA extracted from tissue and blood samples from patients with different cancer types who carried the *KRAS* G12D SNV identified by deep sequencing (Figure 2f). After A-Star processing, the SNV was successfully detected by TaqMan PCR and Sanger sequencing (Figure 2g and Figure S17, Supporting Information). The VAFs of tissue samples processed by A-Star directly increased up to 60-90% from the original VAFs of less than 20%. In particular, the LUAD sample with an extremely low VAF (below 1%) was enriched by more than 50%, representing over a 60-fold improvement. For blood

samples, we performed a preamplification step to increase the initial template amount and then the A-Star procedure. The VAFs increased up to 40-85% from the original VAFs of 2-20%. The above results prove that A-Star offers impressive specificity and sensitivity for analyzing high-background DNA samples for clinical application.

Given that Ago's precise recognition capacity is mediated by gDNAs, we envisioned that A-Star in combination with properly designed multiple gDNAs[15] might be suitable for the multiplex detection of SNVs that may appear in clinical samples (**Figure 3**a**)**. First, we tested duplex enrichment by combining the different SNV allele targets between *KRAS* and *PIK3CA* and a deletion mutation target for *EGFR* with a 1% VAF. The samples were mixed with all the required primers, preoptimized gDNA pairs, and *Pf*Ago in one tube (Figure S18, Supporting Information). The degree of enrichment of all targets analyzed by subsequent TaqMan PCR analysis was of a similar magnitude to the values detected for nonmultiplex A-Star assays (Figure 3b and Figure S19, Supporting Information). Under triplex enrichment in one tube, all three targets were more than 40-fold enriched, as detected by TaqMan PCR, implying orthogonal activity of *Pf*Ago when supplied with multiple gDNA pairs. To mimic clinical samples with a complicated DNA background, we also tested triplex A-Star on cfDNA standards with three available VAFs (0.1%, 1%, and 5%) following a similar multiplex procedure to that above but with an initial pre-PCR amplification step. The quantitative analysis revealed 50-90% of SNV alleles for each of *KRAS* and *PIK3CA*, with a range of 0.1% to 5% SNV alleles, except for 0.1% *PIK3CA* (Figure 3c and Figure S20, Supporting Information). For 0.1% *PIK3CA*, enrichment to 21% was observed, which was 210-fold higher enrichment than in the initial sample and was sufficient for the subsequent analyses. In contrast to the correlation of the enrichment efficiency with the initial SNV alleles for *KRAS* and *PIK3CA*, the enrichment of the deletion mutation

of *EGFR* in the triplex sample achieved a nearly 100% mutation frequency for a wide range of initial mutant alleles. Together, the results indicate that the direction of the orthogonal activity of *Pf*Ago by corresponding gDNAs provides a simple single-tube reaction for multiplex enrichment.

In this study, we developed A-Star, a single-tube *Pf*Ago-directed PCR method that enriches and detects rare SNVs with a high sensitivity, specificity and efficiency. This approach shows marked differences from currently available methods in several important aspects. First, our method uses an Ago endonuclease, which can target any DNA site directed solely by paired gDNAs, doing away with the requirement of the CRISPR system for a suitable PAM/PFS in the target DNA sequence[6]. Second, A-Star possesses an extra-high SNV enrichment efficiency since *Pf*Ago is compatible with PCR, thus combining the advantage of WT elimination with simultaneous SNV amplification in one tube, in contrast with the reported Cas9-based or *Tt*Ago-based two-step enrichment methods[10b,14]. Finally, based on a very easily implemented multigDNA design concept, we successfully demonstrated that the reaction with a single *Pf*Ago allows multiplexing rather than requiring tedious work to select and optimize various orthogonal Cas enzymes[11], offering great potential for simultaneously analyzing multiple instances and classes of somatic mutations. We also noted that there were some differences in the enrichment efficiency for a different rare SNVs, and the rules governing the precise discrimination of the WT and SNV alleles in different gene spectra and the optimal conditions are under investigation in our lab. In conclusion, A-Star provides a powerful tool for the enrichment and detection of rare SNVs in samples such as cfDNA, which might be applied to a much broader spectrum of basic research or clinical situations that require genotyping of rare alleles from complex samples.

Experimental Section

Nucleic Acid Preparation: The ssDNA target, gDNA, and primers were synthesized commercially (Sangon Biotech, China). The 620 bp dsDNA target was synthesized by GenScript (Nanjing, China) in the form of the pET-28a-derived plasmid. The plasmid DNA was extracted using a Plasmid DNA MiniPreps Kit (Generay, China). The 130-160 bp dsDNA target was obtained by amplification of the corresponding plasmid with the primers designed using NCBI Primer-BLAST, with parameters set for amplicon size between 130 and 160 nt, primer melting temperatures between 50°C and 60°C, and primer sizes between 18 and 25 nt. For PCR amplification, 2X PCR Precision^{TM} MasterMix (Abm, Canada) was used. The target dsDNA fragments for the titration experiments were quantified using a PikoGreen dsDNA Quantitative Kit (Life iLab Biotech, China). All of the nucleic acids used in this study are listed in Table S1-S4.

PfAgo Expression and Purification: A codon-optimized version of the *Pf*Ago gene was synthesized by GenScript (Nanjing, China), which was designed as the pET28a-derived plasmid pEX-*Pf*Ago with an N-terminal His-tag. The expression plasmid was transformed into *E. coli* BL21(DE3) cells. A 5 mL seed culture was grown at 37°C in LB medium with 50 μg/mL kanamycin and was subsequently transferred to 1 L of LB in a shaker flask containing 50 μg/mL kanamycin. The cultures were incubated at 37°C until the OD_{600} reached 0.8-1.0, and protein expression was then induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, followed by incubation for 16 h at 20°C. Cells were harvested by centrifugation for 20 min at 6,000 rpm, and the cell pellet was collected for later purification. Cell pellets were resuspended in lysis buffer (20 mM Tris/HCl, 1 M NaCl, pH 8.0) and then disrupted using a High Pressure Homogenizer at 600-800 bar for 3

min (Gefran, Italy). The lysates were centrifuged for 30 min at 12,000 rpm at 4°C, after which the supernatants were subjected to Ni-NTA affinity purification with elution buffer (20 mM Tris/HCl, 1 M NaCl, 200 mM imidazole, pH 8.0). Further gel filtration purification using Superdex 200 (GE Tech, USA) was carried out with elution buffer (20 mM Tris/HCl, 1 M NaCl, pH 8.0). The fractions resulting from gel filtration were analyzed by SDS-PAGE, and fractions containing the protein were flash frozen at −80°C in storage buffer (20 mM Tris- HCl, pH 8.0, 250 mM NaCl, 10% (v/v) glycerol).

DNA Cleavage Assays: Generally, *Pf*Ago-mediated cleavage assays were carried out in reaction buffer (15 mM Tris/HCl pH 8.0, 250 mM NaCl, and 0.5 mM $MnCl_2$)[12d]. For ssDNA cleavage, 0.2 μM *Pf*Ago, 2 μM gDNA, and 0.8 μM ssDNA target were mixed in reaction buffer and then incubated for 15 min at 95°C in a thermocycler (Eppendorf, Germany). Following high-temperature incubation, the samples were cooled down by slowly lowering the temperature at a rate of 0.1°C/s until it reached 10°C. The reactions were stopped via the addition of loading buffer (95% formamide, 0.5 mmol/L EDTA, 0.025% bromophenol blue, 0.025% xylene cyanol FF) at a 1:1 ratio (v/v); the samples were then separated in 16% denaturing polyacrylamide gels and analyzed by staining with GelRed (Biotium, USA). The nucleic acids were visualized using a G:BOX Chemi imager (Syngene, USA), and data were analyzed using Quantity One (Bio-Rad, USA). For the dsDNA cleavage assays, 0.16 μM *Pf*Ago, 2 μM gDNAs, and 58 nM dsDNA target were mixed in reaction buffer before incubation for 15 min at 95°C in a thermocycler (Eppendorf, Germany). Following this high-temperature incubation step, the samples were cooled down by slowly lowering the temperature at a rate of 0.1°C/s until it reached 10°C. The reactions were quenched with 5× DNA loading buffer (Generay, China) and analyzed in 2% agarose gels. The gels were visualized with a Fuji FLA7000 scanner (FUJIFILM Life Science, USA), and the data

were analyzed using Quantity One (Bio-Rad, USA).

PfAgo Activity at Different Temperatures: The effects of temperature on *Pf*Ago activity mediated by different guides were tested across a range of temperatures from 55 to 99°C. Generally, 2.5 μM *Pf*Ago, 2 μM ssDNA guide, and 0.8 μM ssDNA target were mixed in reaction buffer, and the mixture was then incubated for 30 min at 54.8°C, 58.3°C, 60.7°C, 66.0°C, 71.0°C, 74.9°C, 79.9°C, 85.1°C, 90.2°C, 95.1°C, 96.9°C, or 98.7°C. The samples were resolved in 16% denaturing polyacrylamide gels. The gels were stained using GelRed (Biotium, USA). The nucleic acids were visualized using a G:BOX Chemi imager (Syngene, USA), and the data were analyzed using Quantity One (Bio-Rad, USA).

Preparation of Variant Allele Fraction (VAF) Samples: Approximately 160 bp WT and SNV fragments for each of *PIK3CA* E545K, *KRAS* G12D, and *EGFR* delE746-A750 were obtained by PCR amplification. Then, the PCR products were purified using a GeneJET Gel Extraction Kit (Thermo Scientific, USA) and quantified with PikoGreen dsDNA quantitative kits (Life ilab Bio, China). These wild-type and mutant DNA fragments were mixed to obtain VAFs of 1%, 0.1%, 0.01% for each SNV allele in a total DNA concentration of 10 nM for the evaluation of the enrichment performance of A-Star.

A-Star for SNV Enrichment: For the *Pf*Ago-coupled PCR system, 30 nM *Pf*Ago, 600 nM gDNA, 10 nM target DNA with different VAFs, each of forward and reverse primer at 200 nM, and 500 μM Mn^{2+} were mixed in 2X PCR Taq MasterMix (Abm, Canada). Enrichment proceeded in a Mastercycler® RealPlex instrument (Eppendorf, Germany) with a temperature profile of 94°C for 3 minutes, followed by 25 cycles of amplification (94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 20 seconds) and a final 72°C extension for 1 minute. For SNV enrichment in mock cell free-DNA

(cfDNA), mock cfDNA standards simulating actual patient cfDNA samples were purchased from a commercial vendor (Horizon Discovery Group, UK). These standards were provided for each target (*PIK3CA* E545K, *KRAS* G12D, and *EGFR* del) with different VAFs (0.1%, 1%, and 5%) at a concentration of 33.3 ng/µl. Then, 1 µl of these standards was used as input to perform preamplification and subsequently examined using A-Star. The preamplification procedure was carried out in a 25 µl reaction volume using 33.3 ng of mock cfDNA, 2X PCR Precision^{TM} MasterMix (Abm, Canada), and each forward and reverse primer at 250 nM, using a Mastercycler® RealPlex instrument (Eppendorf, Germany) with a temperature profile of 94°C for 3 minutes, followed by 30 cycles of amplification (94°C for 10 seconds, 55°C for 30 seconds, and 72°C for 20 seconds) and a final 72°C, extension for 1 minute. Two microliters of the preamplification products was used as input to perform A-Star. The A-Star system consisted of 20-30 nM *Pf*Ago, gDNAs at a concentration 20-fold that of *Pf*Ago, each forward and reverse primer at 200 nM, and 500 µM Mn^{2+}, which were mixed in 2X PCR Taq MasterMix (Abm, Canada). Enrichment proceeded in a Mastercycler® RealPlex instrument (Eppendorf, Germany) as described above.

For SNV enrichment in tissue samples, 33 ng of extracted genomic DNA from each tissue sample was used as the input to perform preamplification, and the product was then examined using A-Star as described above for the SNV enrichment of mock cfDNA. For SNV enrichment in cfDNA samples, due to the low concentration of extracted cfDNA, 3.3 ng of extracted genomic DNA from each sample was used as the input to perform preamplification, and then 2 µl of preamplification products was used as the input to perform A-Star. The A-Star system consisted of 10-20 nM *Pf*Ago, gDNAs at a concentration 20-fold that of *Pf*Ago, each forward and reverse primer at 200 nM, and 500 µM Mn^{2+}, which were mixed in 2X PCR Taq MasterMix (Abm,

Canada). Enrichment proceeded in a Mastercycler® RealPlex instrument (Eppendorf, Germany) as described above.

Triplex SNV Enrichment of A-Star: Triplex enrichment was performed in 25 µL reaction volumes with 200 nM *Pf*Ago, each gDNA at 4000 nM, 500 µM Mn^{2+}, each forward and reverse primer at 200 nM, 0.4 mM dNTPs (Sangon, China), 0.5 µl of Taq DNA Polymerase (Abm, Canada), and three VAF targets each at a concentration of 10 nM in 10X PCR buffer (Abm, Canada). The reactions were carried out in a Mastercycler® RealPlex instrument (Eppendorf, Germany) with a temperature profile of 94°C for 5 minutes, followed by 25 cycles of amplification (94°C for 30 seconds, 55°C for 30 seconds, 72°C for 20 seconds) and a final 72°C extension for 1 minute.
For triplex SNV enrichment of mock cfDNA, cfDNA standards purchased from a commercial vendor (Horizon Discovery Group, UK) were used as the three targets (*PIK3CA* E545K, *KRAS* G12D, and *EGFR* del) with different VAFs (0.1%, 1%, and 5%) at a concentration of 33.3 ng/µl. Then, 1 µl of these standards was used as the input to perform preamplification, followed by analysis using A-Star. The preamplification process was carried out in a 25 µl reaction volume using 33.3 ng of mock cfDNA, 2X PCR Precision™ MasterMix (Abm, Canada), and each forward and reverse primer at 250 nM for three targets, using a Mastercycler® RealPlex instrument (Eppendorf, Germany) with a temperature profile of 94°C for 3 minutes, followed by 30 cycles of amplification (94°C for 10 seconds, 55°C for 30 seconds, and 72°C for 20 seconds), and a final 72°C, extension for 1 minute. Then, 2 µl of the preamplification product was used as the input to perform A-Star as described above for the SNV enrichment of mock cfDNA.

Detection of Enriched Products by Sanger Sequencing and TaqMan Real-Time PCR: The A-Star enriched products were checked for quality and yield by running 5 µl

of the products in 2.0% agarose gels and visualized on a Fuji FLA7000 scanner (FUJIFILM Life Science, USA), then processed directly for Sanger sequencing (Sangon, China). The primers and probes for the targets (*PIK3CA* E545K, *KRAS* G12D, and *EGFR* del) were designed using Beacon Designer's standard assay design pipeline and ordered as individual primers and probes from Life Technologies (Thermo Fisher, USA). To validate the enriched mutant products, assays were performed using AceQ® qPCR Probe Master Mix (Vazyme biotech co., Ltd), and the results were quantified in a StepOnePlus™ Real-Time PCR System (Thermo Fisher) with the primer and probes sets (Supplementary Data Table S4) at final concentrations of 250 nM primers and 200 nM probes. The temperature profile for amplification consisted of an activation step at 95°C for 8 minutes, followed by 40 cycles of amplification (95°C for 15 seconds and 60°C for 45 seconds).

Collection and DNA extraction of patient samples: Patients with cancers of the pancreas, colorectum, lung, or breast were recruited from Shanghai General Hospital under methods approved by the Human Research Committee of Shanghai General Hospital. The 'healthy donor' samples consisted of peripheral blood samples obtained from 2 individuals with no history of cancer. The cancer and healthy control samples listed in Table S5 were processed in an identical manner. DNA extraction from formalin-fixed, paraffin-embedded (FFPE) tumor blocks was performed using a QIAamp® DNA Blood Mini Kit (Qiagen, Germany) according to the recommended kit protocol. Five milliliters of peripheral venous blood was collected in Streck Cell-Free DNA BCT (Streck Inc., USA), followed by centrifugation at 1,600 g for 10 min at 10°C, and the resulting supernatant was clarified by additional centrifugation. Clarified plasma was transferred to a fresh tube, and DNA was immediately extracted using a QIAamp® DNA Blood Mini Kit (Qiagen, Germany). Finally, exome sequencing and

data processing to produce a BAM file were performed using established NGS analytical pipelines at Shanghai General Hospital to validate the SNV alleles of the samples.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

We thank Yu Ming from Shanghai Jiao Tong University for provided helpful comments in editing this manuscript. This research was supported in part by the grants from Natural Science Foundation of China (31770078) and Ministry of Science and Technology (2017YFE0103300).

Conflict of Interest

Shanghai Jiao Tong University has applied for a patent (China application no. 2019103245802) on A-Star with Y.F., Q.L., G.X., X.G., Z.L., and Y.C. listed as co-inventors.

References

[1] a) C. Bettegowda, M. Sausen, R. J. Leary, I. Kinde, Y. Wang, N. Agrawal, B. R. Bartlett, H. Wang, B. Luber, R. M. Alani, E. S. Antonarakis, N. S. Azad, A. Bardelli, H. Brem, J. L. Cameron, C. C. Lee, L. A. Fecher, G. L. Gallia, P. Gibbs, D. Le, R. L. Giuntoli, M. Goggins, M. D. Hogarty, M. Holdhoff, S. M. Hong, Y. Jiao, H. H. Juhl, J. J. Kim, G. Siravegna, D. A. Laheru, C. Lauricella, M. Lim, E. J. Lipson, S. K. Marie, G. J. Netto, K. S. Oliner, A. Olivi, L. Olsson, G. J. Riggins, A. Sartore-Bianchi, K. Schmidt, M. Shih l, S. M. Oba-Shinjo, S. Siena, D. Theodorescu, J. Tie, T. T. Harkins, S. Veronese, T. L. Wang, J. D. Weingart, C. L. Wolfgang, L. D. Wood, D. Xing, R. H. Hruban, J. Wu, P. J. Allen, C. M. Schmidt, M. A. Choti, V. E. Velculescu, K. W. Kinzler, B. Vogelstein, N. Papadopoulos, L. A. Diaz, Jr., *Sci. Transl. Med.* **2014**, 19, 224; b) C. Abbosh, N. J. Birkbak, G. A. Wilson, M. Jamal-Hanjani, T. Constantin, R. Salari, J. Le Quesne, D. A. Moore, S. Veeriah, R. Rosenthal, T. Marafioti, E. Kirkizlar, T. B. K. Watkins, N. McGranahan, S. Ward, L. Martinson, J. Riley, F. Fraioli, M. Al Bakir, E. Gronroos, F. Zambrana, R. Endozo, W. L. Bi, F. M. Fennessy, N. Sponer, D. Johnson, J. Laycock, S. Shafi, J. Czyzewska-Khan, A. Rowan, T. Chambers, N.

Matthews, S. Turajlic, C. Hiley, S. M. Lee, M. D. Forster, T. Ahmad, M. Falzon, E. Borg, D. Lawrence, M. Hayward, S. Kolvekar, N. Panagiotopoulos, S. M. Janes, R. Thakrar, A. Ahmed, F. Blackhall, Y. Summers, D. Hafez, A. Naik, A. Ganguly, S. Kareht, R. Shah, L. Joseph, A. Marie Quinn, P. A. Crosbie, B. Naidu, G. Middleton, G. Langman, S. Trotter, M. Nicolson, H. Remmen, K. Kerr, M. Chetty, L. Gomersall, D. A. Fennell, A. Nakas, S. Rathinam, G. Anand, S. Khan, P. Russell, V. Ezhil, B. Ismail, M. Irvin-Sellers, V. Prakash, J. F. Lester, M. Kornaszewska, R. Attanoos, H. Adams, H. Davies, D. Oukrif, A. U. Akarca, J. A. Hartley, H. L. Lowe, S. Lock, N. Iles, H. Bell, Y. Ngai, G. Elgar, Z. Szallasi, R. F. Schwarz, J. Herrero, A. Stewart, S. A. Quezada, K. S. Peggs, P. Van Loo, C. Dive, C. J. Lin, M. Rabinowitz, H. Aerts, A. Hackshaw, J. A. Shaw, B. G. Zimmermann, T. R. consortium, P. consortium, C. Swanton, *Nature* **2017**, 545, 446.

[2] a) J. J. Salk, M. W. Schmitt, L. A. Loeb, *Nat. Rev. Genet.* **2018**, 19, 269; b) C. A. Milbury, J. Li, G. M. Makrigiorgos, *Clin. Chem*. **2009**, 55, 632.

[3] a) C. R. Newton, A. Graham, L. E. Heptinstall, S. J. Powell, C. Summers, N. Kalsheker, J. C. Smith, A. F. Markham, *Nucleic. Acids. Res.* **1989**, 17, 2503; b) D. Y. Vargas, F. R. Kramer, S. Tyagi, S. A. Marras, *PLoS. One*. **2016**, 11, e0156546; c) L. R. Wu, S. X. Chen, Y. Wu, A. A. Patel, D. Y. Zhang, *Nat. Biomed. Eng*. **2017**, 1, 714.

[4] a) M. Urdea, L. A. Penny, S. S. Olmsted, M. Y. Giovanni, P. Kaspar, A. Shepherd, P. Wilson, C. A. Dahl, S. Buchsbaum, G. Moeller, D. C. Hay Burgess, *Nature* **2006**, 444 Suppl 1, 73; b) E. Crowley, F. Di Nicolantonio, F. Loupakis, A. Bardelli, Nat. Rev. Clin. Oncol. **2013**, 10, 472; c) E. Heitzer, I. S. Haque, C. E. S. Roberts, M. R. Speicher, *Nat. Rev. Genet*. **2019**, 20, 71.

[5] a) C. Song, Y. Liu, R. Fontana, A. Makrigiorgos, H. Mamon, M. H. Kulke, G. M. Makrigiorgos, *Nucleic. Acids. Res.* **2016**, 44, e146; b) L. Xiao, D. X. Luo, W. Pu, W. Luo, Y. Pan, H. Xu, W. Li, R. Zhang, H. Wang, F. Wang, X. Liu, Y. Sun, D. F. Liao, Y. Feng, C. Xing, P. Sirois, J. Zhang, N. He, K. Li, *J. Biomed. Nanotechnol*. **2019**, 15, 1052.

[6] a) C. H. Huang, K. C. Lee, J. A. Doudna, *Trends. Cancer.* **2018**, 4, 499; b) Y. Li, S. Li, J. Wang, G. Liu, *Trends. Biotechnol.* **2019**, 37, 730; c) C. Chiu, *Cell. Host. Microbe.* **2018**, 23, 702.

[7] a) M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, *Science* **2012**, 337, 816; b) C. Anders, O. Niewoehner, A. Duerst, M. Jinek, *Nature* **2014**, 513, 569.

[8] a) O. O. Abudayyeh, J. S. Gootenberg, S. Konermann, J. Joung, I. M. Slaymaker, D. B. Cox, S. Shmakov, K. S. Makarova, E. Semenova, L. Minakhin, K. Severinov, A. Regev, E. S. Lander, E. V. Koonin, F. Zhang, *Science* **2016**, 353, aaf5573; b) A. East-Seletsky, M. R. O'Connell, S. C. Knight, D. Burstein, J. H. Cate, R. Tjian, J. A. Doudna, *Nature* **2016**, 538, 270.

[9] a) J. S. Gootenberg, O. O. Abudayyeh, J. W. Lee, P. Essletzbichler, A. J. Dy, J. Joung, V. Verdine, N. Donghia, N. M. Daringer, C. A. Freije, C. Myhrvold, R. P. Bhattacharyya, J. Livny, A. Regev, E. V. Koonin, D. T. Hung, P. C. Sabeti, J. J. Collins, F. Zhang, *Science* **2017**, 356, 438; b) J. S. Chen, E. Ma, L. B. Harrington, M. Da Costa, X. Tian, J. M. Palefsky, J. A. Doudna, *Science* **2018**, 360, 436; c) S. Y. Li, Q. X. Cheng, J. M. Wang, X. Y. Li, Z. L. Zhang, S. Gao, R. B. Cao, G. P. Zhao, J. Wang, *Cell. Discov.* **2018**, 4, 20.

[10] S. H. Lee, J. Yu, G. H. Hwang, S. Kim, H. S. Kim, S. Ye, K. Kim, J. Park, D. Y. Park, Y. K. Cho, J. S. Kim, S. Bae, *Oncogene* **2017**, 36, 6823.

[11] J. S. Gootenberg, O. O. Abudayyeh, M. J. Kellner, J. Joung, J. J. Collins, F. Zhang, *Science* **2018**, 360, 439.

[12] a) D. C. Swarts, K. Makarova, Y. Wang, K. Nakanishi, R. F. Ketting, E. V. Koonin, D. J. Patel, J. van der Oost, *Nat. Struct. Mol. Biol*. **2014**, 21, 743; b) J. W. Hegge, D. C. Swarts, J. van der Oost, *Nat. Rev. Microbiol.* **2018**, 16, 5; c) Y. Wang, S. Juranek, H. Li, G. Sheng, T. Tuschl, D. J. Patel, *Nature* **2008**, 456, 921; d) D. C. Swarts, J. W. Hegge, I. Hinojo, M. Shiimori, M. A. Ellis, J. Dumrongkulraksa, R. M. Terns, M. P. Terns, J. van der Oost, *Nucleic. Acids. Res.* **2015**, 43, 5120; e) E. Kaya, K. W. Doxzen, K. R. Knoll, R. C. Wilson, S. C. Strutt, P. J. Kranzusch, J. A. Doudna, *Proc Natl Acad Sci U S A.* **2016**, 113, 4057; f) S. Willkomm, C. A. Oellig, A. Zander, T. Restle, R. Keegan, D. Grohmann, S. Schneider, *Nat. Microbiol.* **2017**, 2, 17035.

[13]A. Lapinaite, J. A. Doudna, J. H. D. Cate, *Proc Natl Acad Sci U S A.* **2018**, 115, 3368.

[14] J. Song, J. W. Hegge, M. G. Mauk, N. Bhagwat, J. E. Till, L. T. Azink, J. Peng, M. Sen, J. Chen, J. Mays, E. Carpenter, J. van der Oost, H. H. Bau, *BioRxiv* [preprint] **2018**, Dec 10.

[15]B. Enghiad, H. Zhao, *ACS Synth. Biol.* **2017**, 6, 752.

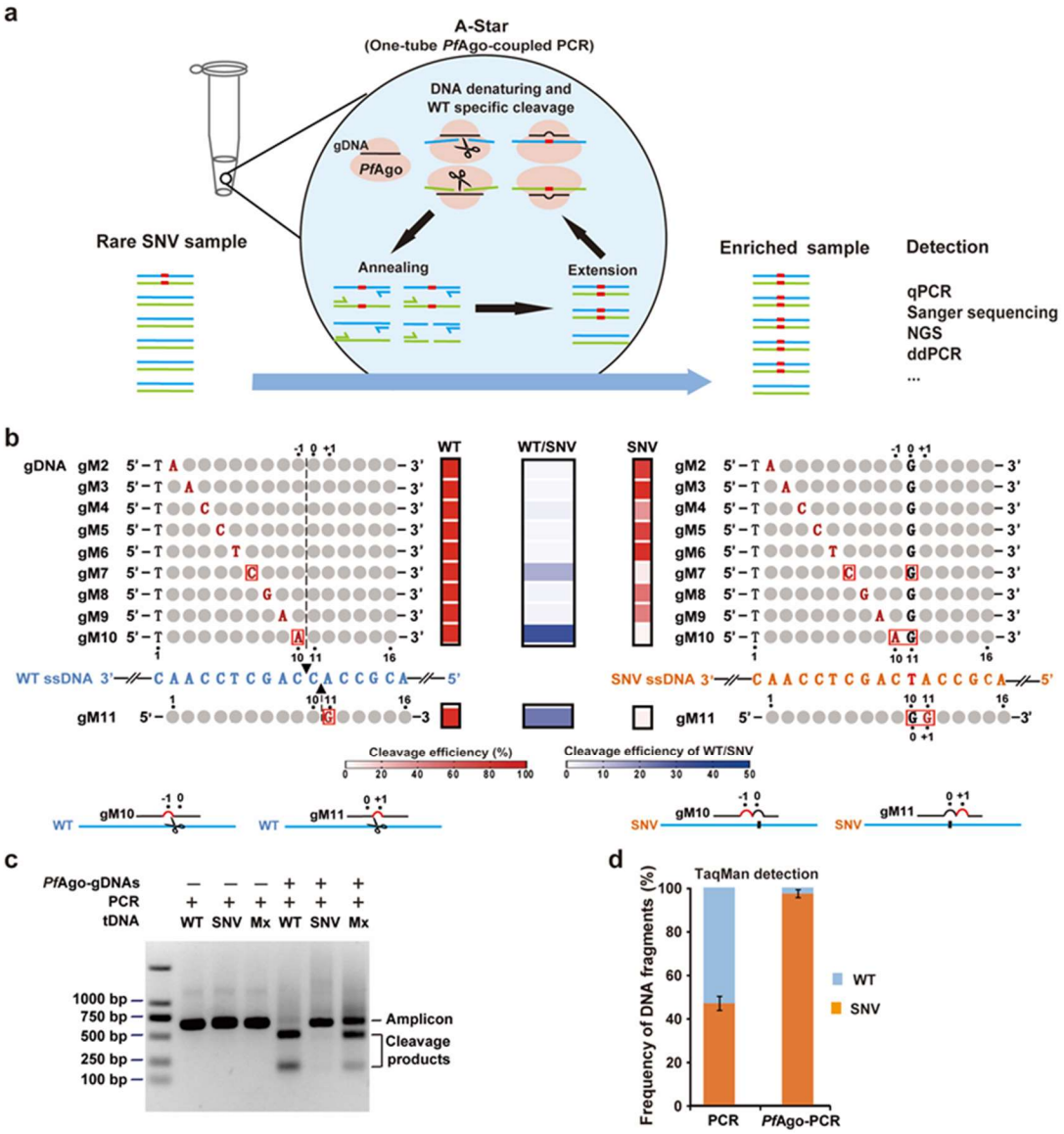

Figure 1. A-Star enables *Pf*Ago-mediated discriminating cleavage for efficient SNV enrichment and detection. a) Schematic representation of the A-Star approach for rare SNV enrichment. Thermophilic *Pf*Ago-mediated cleavage of the wild-type sequence occurs during the denaturing step at 95°C, and the retained SNV is then used as a template for subsequent PCR amplification cycles, resulting in the efficient enrichment and detection of rare SNVs. b) Schematic representation of gDNAs designed to guide the specific cleavage of the *KRAS* G12D ssDNA target. A mismatched nucleotide was introduced to each gDNA between positions 2 and 11; the mismatched nucleotides are highlighted in red, and the SNV site is shown in black. c) Gel electrophoresis showed

that *Pf*Ago-gDNA can be coupled with PCR in a single reaction to cleave WT sequences, thereby enriching rare SNVs. gDNA: guide DNA; tDNA: target DNA. d) Evaluation of the enrichment results by TaqMan PCR for *Pf*Ago compatible with PCR to efficiently enrich SNVs. A sample mixture composed of *KRAS* WT and G12D SNV fragments at equal molar ratios was used as the input for either PCR or *Pf*Ago-coupled PCR. Error bars represent the mean ± s.d., n = 3.

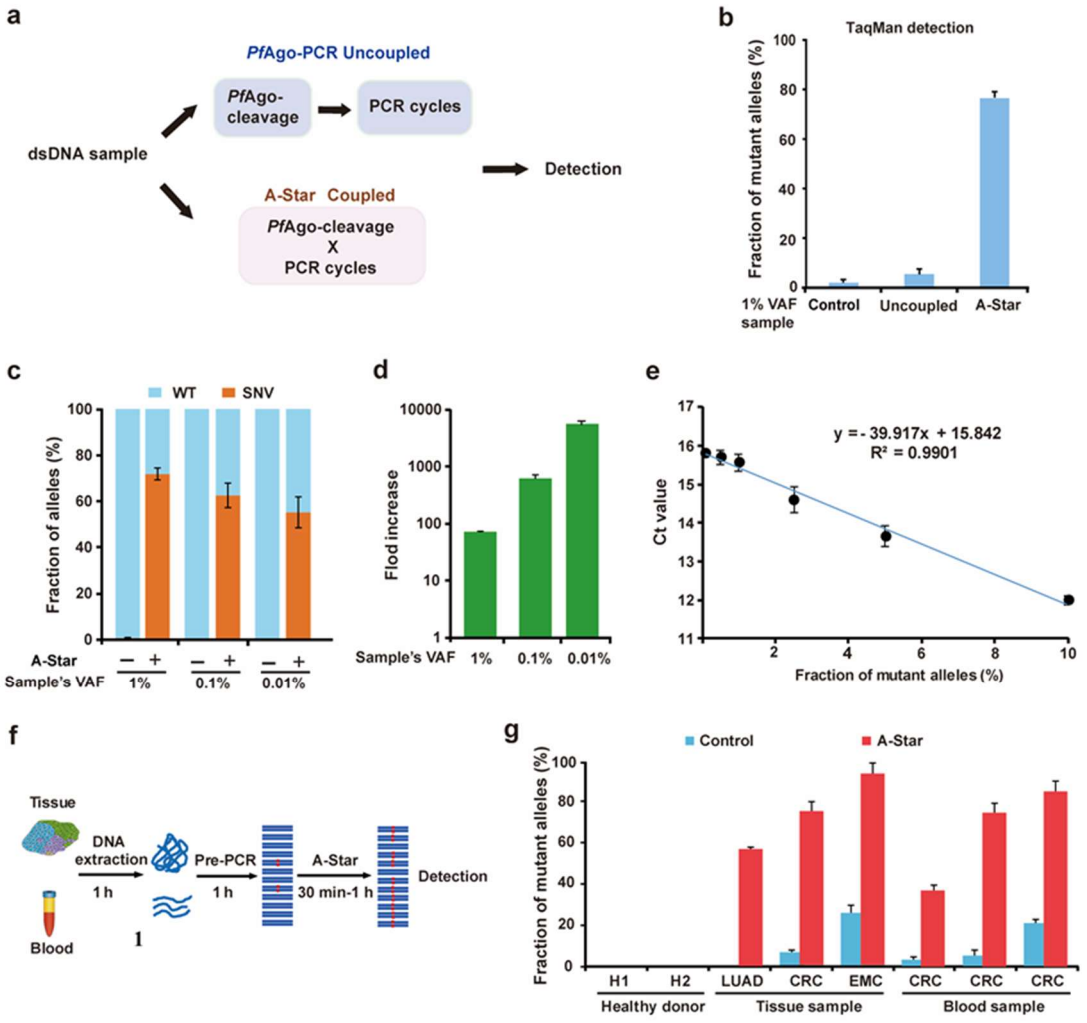

Figure 2. Highly sensitive enrichment and rapid identification of a rare SNV by A-Star for test templates as well as clinical samples. a) Schematic outline of rare SNV enrichment by either single-tube A-Star or the uncoupled *Pf*Ago-PCR approach. b) A-Star can efficiently enrich a 1% VAF sample of *KRAS* G12D compared with the uncoupled *Pf*Ago-PCR approach according to TaqMan PCR. The control reactions did not contain gDNA pairs. c, d) *KRAS* G12D samples with varying VAFs were enriched by A-Star and quantified using TaqMan PCR to support efficient A-Star enrichment at a 0.01% VAF (c) by over 6000-fold (d). e) TaqMan PCR assay showing that the threshold cycle (Ct) value presents a linear correlation with the fraction of mutant alleles. f) Schematic representation of the A-Star procedure for the identification of cancer-associated SNVs in human samples. g) A-Star is capable of evaluating clinical

samples from diverse cancer types, as exemplified by *KRAS* G12D-containing samples from patients with lung adenocarcinoma (LUAD), colorectal cancer (CRC), or endometrial cancer (EMC). The control reactions did not contain gDNA pairs. Error bars represent the mean ± s.d., n = 3.

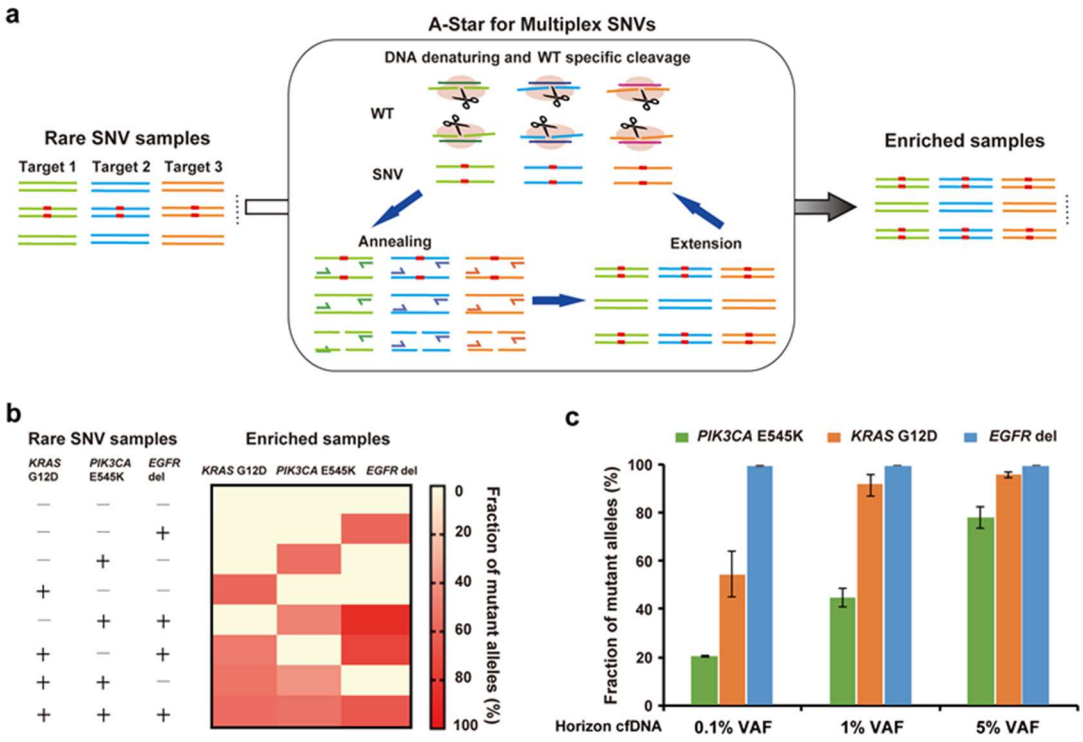

Figure 3. Multiplex A-Star can detect multiple SNVs in the same reaction due to the orthogonal cleavage activity of *Pf*Ago. a) Schematic representation of multiplex SNV detection in a one-pot reaction containing multiple pairs of primers and gDNAs. b) Multiplex detection of *KRAS* G12D, *PIK3CA* E545K, and *EGFR* del by A-Star, showing the orthogonal cleavage activity of *Pf*Ago directed by the corresponding gDNAs. Different combinations of the three SNV alleles were tested by A-Star using the same reaction conditions, including *Pf*Ago and three pairs of gDNAs. c) A-Star triplex enrichment of *KRAS* G12D, *PIK3CA* E545K, and *EGFR* del using standard cfDNA samples with different VAFs of 0.1%, 1%, and 5%. Solutions containing 33 ng/μL of these standards were used as the input, and the output was detected using TaqMan PCR. The control reactions did not contain gDNA pairs. Error bars represent the mean ± s.d., n = 3 for each SNV.