www.nature.com/mtna

Application of Peptide Nucleic Acid-based Assays Toward Detection of Somatic Mosaicism

Christopher S Hong¹, Chunzhang Yang² and Zhengping Zhuang¹

Peptide nucleic acids (PNAs) are synthetic oligonucleotides with many applications. Compared with DNA, PNAs bind their complementary DNA strand with higher specificity and strength, an attribute that can make it an effective polymerase chain reaction clamp. A growing body of work has demonstrated the utility of PNAs in detecting low levels of mutant DNA, particularly in the detection of circulating mutated tumor cells in the peripheral blood. The PNA-based assay has greater sensitivity than direct sequencing and is significantly more affordable and rapid than next-generation deep sequencing. We have previously demonstrated that PNAs can successfully detect somatic mosaicism in patients with suspected disease phenotypes. In this report, we detail our methodology behind PNA design and application. We describe our protocol for optimizing the PNA for sequencing use and for determining the sensitivity of the PNA-based assay. Lastly, we discuss the potential applications of our assay for future laboratory and clinical purposes and highlight the role of PNAs in the detection of somatic mosaicism. *Molecular Therapy—Nucleic Acids* (2016) **5**, e314; doi:10.1038/mtna.2016.22; published online 26 April 2016

Subject Category: Nucleic Acid Chemistries

Introduction

Somatic mosaicism is defined as the presence of two or more genotypes in somatic cells of an individual who was conceived from a single fertilized egg. While uncorrected genetic alterations are thought to occur approximately every two mitotic divisions, the vast majority of these mutations are either deleted through programmed apoptosis or immune surveillance.^{1,2} In most cases, mutations that persist beyond this initial checkpoint do not exert any pathologic significance or fail to accumulate to a degree necessary for disease manifestation. However, mutations occurring early in life, especially in the immediate postzygotic period, may provide the genetic predisposing material for later disease development, particularly in regards to cancer.³ Well-established examples of this phenomenon include McCune-Albright syndrome and tuberous sclerosis, predisposing to development of polyostotic fibrous dysplasias and hamartomas, respectively.^{4,5} In addition, diseases classically associated with autosomal inheritance patterns have also been shown to occur sporadically in the context of somatic mosaicism, such as Neurofibromatosis type 1, Neurofibromatosis type 2, and Von-Hippel-Lindau syndrome.^{5–7} In recent years, there has been a resurgence of interest in the role of somatic mosaicism in nonsyndromic cases of tumorigenesis and noncancerous human disease.8-10

Originally developed by Frederick Sanger and colleagues in 1977, Sanger sequencing has become one of the most widely used methods of DNA sequencing in clinical and basic research. The method utilizes selective incorporation of chain-terminating dideoxynucleotides to sequence predetermined areas of the genome and is both affordable and rapid. However, the sensitivity of this method is limited, particularly for the detection of somatic mosaicism, with a detection threshold of ~15–20%.¹¹ In addition, Sanger sequencing of mosaicism present in greater than 60–70% of cells can be mistaken for germline variants.¹² More sensitive sequencing methods such as deep-coverage, "next-generation" sequencing, although increasingly accessible, are still time-consuming and costly.¹³

Peptide nucleic acids (PNAs) are synthetic oligonucleotides, in which the native sugar-phosphate backbone of DNA has been replaced with amino acids.¹⁴ Compared with DNA, PNAs bind their complementary sequences with greater affinity and specificity, due to high resistance to nuclease and protease-mediated degradation. In addition to other diagnostic and therapeutic applications, PNAs represent an affordable and rapid means of testing for somatic mosaicism, by binding to and inhibiting polymerase chain reaction (PCR) amplification of a sequence of interest.

In this report, we argue for the utility of PNAs in the discovery of somatic mosaicism. We describe the design and optimization of PNAs for diagnostic sequencing and discuss recent work from our group, demonstrating the efficacy of PNAs in detecting somatic mosaicism. Lastly, we discuss the advantages and disadvantages of PNA technology and suggest that PNA-based sequencing may be a highly affordable and rapid means of detecting somatic mosaicism.

Results

PNA design

PNAs are synthetic oligonucleotides that share properties of both peptides and nucleic acids. Structurally, the PNA backbone consists of repeating *N*-(2-aminoethyl)-glycine units linked by amide bonds, to which purine and pyrimidine bases are attached. Unlike native DNA, it lacks the pentose sugar moieties or phosphate groups. In addition, unlike DNA, the

Keywords: gene sequencing; oligonucleotide; peptide nucleic acid; polymerase chain reaction; somatic mosaicism Received 2 January 2016; accepted 6 February 2016; published online 26 April 2016. doi:10.1038/mtna.2016.22

¹Surgical Neurology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland, USA.; ²Neuro-Oncology Branch, Center for Cancer Research, National Cancer Institute, Bethesda, Maryland, USA. Correspondence: Zhengping Zhuang, 10 Center Drive, Room 3D20, Bethesda, MD 20892, USA. E-mail: zhuangp@ninds.nih.gov

PNA backbone is acyclic, achiral, and importantly, neutral, which contributes to higher specificity and stronger binding between complementary PNA/DNA strands than DNA/DNA strands. These properties also cause higher melting temperatures, and thus, PNA hybridization requires higher than normal temperatures during PCR assays. PCR conditions are explained in greater detail below. A deeper discussion of PNA design is outside the scope of this manuscript; however, Nielsen and Egholm,^{14,15} who contributed to the original design of the PNA, have explained their work in-depth.

PCR conditions

In our experience, we perform PNA-PCR assays with the following reagents: 200–500 ng genomic DNA, 25 μ I Taq 2x Master Mix (New England BioLabs, Ipswich, MA, USA), 0.25 μ I each of forward and reverse primers (100 μ mol/l), and 0.5 μ I PNA (in the optimal concentration; explained in greater detail in the next section) in a final volume of 50 μ I. The PCR conditions we use were initially described by Lietman *et al.*¹⁶ and are as follows: denaturation at 94 °C for 15 minutes, followed by 40 cycles of: 94 °C for 30 seconds (denaturation), 68 °C for 60 seconds (PNA hybridization), 55 °C for 30 seconds (primer annealing), and 72 °C for 7 minutes.

PNA optimization

The PNA is typically reconstituted into a 100 nmol/l solution for ease of storage. However, in our experience, this results in too high of a concentration of the PNA and completely abrogates amplification of genomic DNA. To determine the upper limit of optimal PNA concentration, 200–500 ng of wild-type genomic DNA is subjected to PCR under PNA-PCR conditions with serially diluted concentrations of the PNA. We typically dilute by 10-fold (1:10, 1:100, 1:1,000, 1:10,000) and later optimize with smaller increments. Subsequently, the PCR products are run under gel electrophoresis, and gels are analyzed under ultraviolet light. The band intensity of the electrophoresed products diminishes as the concentration of the PNA increases, and the PNA concentration at which little to no band is seen is determined to be the upper limit of optimal PNA concentration (**Figure 1a**). We have found that the optimal PNA concentration can vary greatly. For instance, in the aforementioned study examples, the concentrations used for the HIF2A and IDH1 PNAs were 2×10^{-14} mol (1:3,000) and 1×10^{-12} mol (1:1,000), respectively.

PNA application

Once the optimal PNA concentration is determined, the PNA-PCR assay can be applied to the sample DNA of interest, and the reaction products are run under gel electrophoresis. The band is then excised, purified utilizing a standard gel purification kit (MinElute Gel Extraction Kit; Qiagen, Venlo, Netherlands), and sent for routine Sanger sequencing. In the case of somatic mosaicism, a classic double peak or amplified single mutant allele at the nucleotide site may be seen. In instances of absent somatic mosaicism, persistent wild-type sequence is typically observed (Figure 1b).

PNA sensitivity

The sensitivity of the PNA is determined by performing the PNA-PCR assay on plasmid mixtures comprised of varying proportions of mutant and wild-type plasmids. The methods are further detailed below.

Figure 1 Optimization and application of IDH1 PNA. (a) Gel electrophoresis of PCR products derived from IDH1 (codon 132) sequencing of wild-type DNA with serial dilutions of 100 nmol/l IDH1-specific PNA. The far left column is a 1-kb DNA ladder. Subsequently, from left to right are 1,000-fold (1/1,000), 250-fold (1/250), etc. dilutions of 100 nmol/l PNA. The band intensity is significantly reduced from 1/1,000 to 1/250 and is completely absent at 1/62.5, suggesting complete blockage of amplification by the PNA. A dilution of 1/1,000 was chosen as the appropriate dilution for this IDH1-specific PNA. (b) Sanger sequencing of IDH1 (R132) from leukocyte DNA is shown with and without PNA. Direct sequencing (top panel) demonstrated wild-type IDH1 R132 status. However, addition of the IDH1-specific PNA revealed an IDH1

(R132H) mutation, suggestive of IDH1 somatic mosaicism. PNA, peptide nucleic acids.





First, mutant and wild-type plasmids are generated through thymine adenine (TA) cloning technology. We utilize heterozygous PCR products from standard PCR of known mutated DNA (*i.e.*, tumor tissue) and clone them into pCR4-TOPO vectors using the TOPO TA Cloning Kit for Sequencing (Invitrogen, Carlsbad, CA). After confirmation of wild-type and mutant DNA sequences of the inserted DNA fragments, the plasmids are amplified with PCR and used for further PCR templates.

Subsequently, wild-type (or mutant) plasmids are diluted in various concentrations, to determine the minimum concentration of plasmids required for the PNA-PCR assay to generate a detectable band on gel electrophoresis. We typically utilize plasmid concentrations ranging from 10 copies/reaction to 10^{10} copies/reaction. In our experience, we have found that ~ 10^6 – 10^7 copies/reaction to be the minimum copy number required for gel electrophoresis detection (**Figure 2a**). Next, varying mixtures of mutant and wild-type plasmids are created in which the total number of plasmids per reaction equals the minimum copy number predetermined for gel electrophoresis detection (*i.e.*, 10^7

copies). We typically create mixtures utilizing the following percentages of mutant plasmid DNA in each reaction: 90, 80, 70, 60, 50, 40, 30, 20, 10, 5, 2.5, 1.25, 0.6125, and 0.30625%. Next, 1 µl of each plasmid mixture is subjected to the PCR-PNA assay, and all visible bands on gel electrophoresis are then excised, purified, and sent for Sanger sequencing (Figure 2b). Utilizing open-access sequencing analysis software (*i.e.*, 4Peaks software; Nucleobytes, Aalsmeer, Netherlands), mutant to wild-type signal ratios measured from each double peak are plotted to generate exponential curves. Lastly, the same experiment is performed with the plasmid mixtures in a standard PCR reaction without the PNA, to generate a control curve. The percentage of mutant plasmid DNA that corresponds to the lowest detectable mutant allele signal on the curve is determined to be the lower limit for sensitivity of the PNA (Figure 2c).

Study examples

In this section, we briefly describe two studies recently reported by our group that demonstrate the efficacy of PNAs.







In 2012, we described a novel syndrome of paraganglioma, somatostatinoma, and congenital polycythemia secondary to tumor-specific mutations in hypoxia-inducible factor 2A (HIF2A).¹⁷ At the time, we had hypothesized that the pathophysiology underlying this disorder was somatic mosaicism of HIF2A mutations. However, through standard Sanger sequencing, we were unable to detect HIF2A mutations in germ-line DNA of the two patients we described. Recently, we developed a HIF2A-specific PNA targeting the mutation site present in the patients' tumors. With this PNA, we discovered the presence of the same tumor-specific HIF2A mutation in leukocytederived DNA, as well as DNA extracted from hair, nail, and saliva.¹⁸ As such, PNA technology enabled us to validate our initial hypothesis that somatic mosaicism of HIF2A mutations was the genetic mechanism underlying the syndrome of paraganglioma, somatostatinoma, and congenital polycythemia.

In a separate study, we investigated Maffucci syndrome. a disorder characterized by enchondromatosis and benign hemangioma formation. In 2011, Amary and colleagues¹⁹ showed that patients with this condition harbored somatic mosaicism of mutations in isocitrate dehvdrogenase 1 (IDH1) or IDH2. However, other tumors have been described in these patients, including those of the central nervous system. Recently, we described a patient with Maffucci syndrome who developed a pituitary adenoma as well as a chondrosarcoma of the skull base.20 While we were able to detect an IDH1 (R132C) mutation in the chondrosarcoma with standard sequencing techniques, we could not detect the presence of the same IDH1 mutation in tissue from the pituitary adenoma, which was extremely limited and diluted amonast a significant background of nontumoral cells. However, with the aid of an IDH1-specific PNA, we detected the same IDH1 (R132C) mutation in the pituitary adenoma and thus provided genetic evidence that pituitary adenomas could indeed arise within the context of somatic mosaicism of IDH mutations in Maffucci syndrome.

Discussion

In this report, we have demonstrated the efficacy of a PNA-PCR assay to detect low levels of somatic mosaicism in germ-line tissues as well as mutations in cell-heterogeneous samples. Furthermore, we have detailed the methodology utilized by our group to design, optimize, apply, and quantify PNA technology for gene sequencing. Other applications of PNAs are many, including PNA-based probes for fluorescence in situ hybridization, PNA-based transcriptomic studies and single-nucleotide polymorphism (SNP) arrays, and therapeutic inhibition of mRNA/ micro-RNAs.^{21,22} Besides our own reports described above, Lietman and colleagues¹⁶ have utilized a PNA-PCR assay to detect somatic mosaicism of mutant quanine nucleotide binding protein. alpha stimulating (GNAS) alleles circulating in the peripheral blood of patients with McCune-Albright syndrome and fibrous dysplasia. Bushman and colleagues²³ recently applied PNAbased probes targeting amyloid precursor protein to demonstrate somatic mosaicism of amyloid precursor protein in single neurons found in sporadic Alzheimer's disease. Further studies may demonstrate similar applications for PNAs in common diseases of somatic mosaic etiology, including the neurofibromatoses, Von-Hippel–Lindau syndrome, and tuberous sclerosis.

Recently, PNA application in genomic sequencing has extended beyond the diagnostic realm toward prognostic utility. In tumor patients with known mutations, the detection of mutated tumor DNA in the bloodstream may predict long-term prognosis as well as response to therapy. Gonzalez-Cao and colleagues²⁴ recently showed that a PNA-based assay detecting the BRAF V600E mutation in circulating free DNA had both prognostic and predictive value in assessing response to BRAF inhibitor treatment. Patients with detectable BRAF V600E in their circulation prior to treatment exhibited significantly worsened progression-free and overall survivals, even after targeted BRAF therapy, Likewise, others have shown that PNA-mediated detection of EGFR L858R (exon 21) mutations in the bloodstream of patients with advanced non small-cell lung cancer predicted poorer outcomes, despite treatment with the tyrosine kinase inhibitor, erlotinib.25

Like our studies, others have demonstrated the increased sensitivity of a combined PNA-PCR assay, compared with traditional Sanger sequencing.25-28 In addition, the sensitivity of PNA-based sequencing compares favorably to next-generation deep sequencing, a gold standard for the detection of low levels of mutant DNA, and may generate higher vield use when used together.^{29,30} Advantages of the PNA application over deep sequencing include significantly reduced costs, which can equal that of standard Sanger sequencing. In addition, unlike many deep sequencing platforms that require prolonged time for sample analysis, results from PNA-PCR assays can be available within one day, making it suitable for routine clinical use. A major disadvantage of PNA-based sequencing remains the inability to detect novel mutations. PNAs must be designed to target known hotspot mutation sites in the genome, although large-scale PNA-based arrays can be designed to permit onetime testing for multiple genes of interest.³¹ Likewise, another limitation of our PNA-PCR assay is the inability to precisely quantify the proportion of mutated cells in a given sample. Although we have previously quantified the degree of somatic mosaicism with TA cloning methods, we have found that a TA cloning-based approach does not work for all samples, particularly those with extremely low levels of mutant DNA.

In conclusion, it is clear from our experience and others that the application of PNA technology toward direct Sanger sequencing yields improved detection of low levels of mutant DNA. In particular, PNAs hold great promise for the detection of somatic mosaicism and understanding disease pathogenesis. Although PNAs remain relatively underutilized, they are a powerful tool and should continue to be incorporated into both basic and clinical research settings.

Materials and Methods

PNAs are commercially available from many different manufacturers. Our group has designed and purchased PNAs from PNA Bio (Thousand Oaks, CA). Other sources that offer custom PNAs include ATDBio (Southampton, UK), Link Technologies (Lanarkshire, Scotland), NeoScientific (Cambridge, MA), and Panagene (Daejeon, Korea).

Acknowledgments This study was supported by the Intramural Research Program of the National Institute of Neurological Disorders and Stroke (NINDS) at the National Institutes of Health (NIH).

- Behjati, S, Huch, M, van Boxtel, R, Karthaus, W, Wedge, DC, Tamuri, AU *et al.* (2014). Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* 513: 422–425.
- Seshadri, R, Kutlaca, RJ, Trainor, K, Matthews, C and Morley, AA (1987). Mutation rate of normal and malignant human lymphocytes. *Cancer Res* 47: 407–409.
- Erickson, RP (2010). Somatic gene mutation and human disease other than cancer: an update. *Mutat Res* 705: 96–106.
- Boyce, AM and Collins, MT (1993). Fibrous Dysplasia/McCune-Albright Syndrome. In: Pagon, RA et al. (eds.) GeneReviews(R). University of Washington, Seattle, WA, USA.
- Murgia, A, Martella, M, Vinanzi, C, Polli, R, Perilongo, G and Opocher, G (2000). Somatic mosaicism in von Hippel–Lindau Disease. *Hum Mutat* 15: 114.
- Ainsworth, PJ, Chakraborty, PK and Weksberg, R (1997). Example of somatic mosaicism in a series of de novo neurofibromatosis type 1 cases due to a maternally derived deletion. *Hum Mutat* 9: 452–457.
- Evans, DG, Wallace, AJ, Wu, CL, Trueman, L, Ramsden, RT and Strachan, T (1998). Somatic mosaicism: a common cause of classic disease in tumor-prone syndromes? Lessons from type 2 neurofibromatosis. *Am J Hum Genet* 63: 727–736.
- Frank, SA (2010). Evolution in health and medicine Sackler colloquium: somatic evolutionary genomics: mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proc Natl Acad Sci USA* **107 (suppl. 1)**: 1725–1730
- Poduri, A, Evrony, GD, Cai, X and Walsh, CA (2013). Somatic mutation, genomic variation, and neurological disease. *Science* 341: 1237758.
- Youssoufian, H and Pyeritz, RE (2002). Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet* 3: 748–758.
- Rohlin, A, Wernersson, J, Engwall, Y, Wiklund, L, Björk, J and Nordling, M (2009). Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum Mutat* 30: 1012–1020.
- Jamuar, SS, Lam, AT, Kircher, M, D'Gama, AM, Wang, J, Barry, BJ et al. (2014). Somatic mutations in cerebral cortical malformations. N Engl J Med 371: 733–743.
- Rehm, HL, Bale, SJ, Bayrak-Toydemir, P, Berg, JS, Brown, KK, Deignan, JL *et al.*; Working Group of the American College of Medical Genetics and Genomics Laboratory Quality Assurance Commitee (2013). ACMG clinical laboratory standards for next-generation sequencing. *Genet Med* 15: 733–747.
- Buchardt, O, Egholm, M, Berg, RH and Nielsen, PE (1993). Peptide nucleic acids and their potential applications in biotechnology. *Trends Biotechnol* 11: 384–386.
- Nielsen, PE and Egholm, M (1999). An introduction to peptide nucleic acid. Curr Issues Mol Biol 1: 89–104.
- Lietman, SA, Ding, C and Levine, MA (2005). A highly sensitive polymerase chain reaction method detects activating mutations of the GNAS gene in peripheral blood cells in McCune–Albright syndrome or isolated fibrous dysplasia. J Bone Joint Surg Am 87: 2489–2494.
- Zhuang, Z, Yang, C, Lorenzo, F, Merino, M, Fojo, T, Kebebew, E *et al.* (2012). Somatic HIF2A gain-of-function mutations in paraganglioma with polycythemia. *N Engl J Med* 367: 922–930.
- Yang, CH, Hong, CS, Prchal, JT, Balint, MT, Pacak, K, Zhuang, Z (2015). Somatic mosaicism of EPAS1 mutations in the syndrome of paraganglioma and somatostatinoma associated with polycythemia. *Hum Genome Var* (15053): 1–2
- Amary, MF, Damato, S, Halai, D, Eskandarpour, M, Berisha, F, Bonar, F *et al.* (2011). Ollier disease and Maffucci syndrome are caused by somatic mosaic mutations of IDH1 and IDH2. *Nat Genet* 43: 1262–1265.

- Hao, S, Hong, CS, Feng, J, Yang, C, Chittiboina, P, Zhang, J et al. (2015). Somatic IDH1 mutation in a pituitary adenoma of a patient with Maffucci syndrome. J Neurosurg (epub ahead of print). doi: 10.3171/2015.4.JNS15191.
- Gambari, R (2014). Peptide nucleic acids: a review on recent patents and technology transfer. Expert Opin Ther Pat 24: 267–294.
- O'Hara, SD, Stehle, T and Garcea, R (2014). Glycan receptors of the Polyomaviridae: structure, function, and pathogenesis. *Curr Opin Virol* 7: 73–78.
- Bushman, DM, Kaeser, GE, Siddoway, B, Westra, JW, Rivera, RR, Rehen, SK et al. (2015). Genomic mosaicism with increased amyloid precursor protein (APP) gene copy number in single neurons from sporadic Alzheimer's disease brains. *Elife* 4 (e05116): 1–26.
- Gonzalez-Cao, M, Mayo-de-Las-Casas, C, Molina-Vila, MA, De Mattos-Arruda, L, Muñoz-Couselo, E, Manzano, JL *et al.* (2015). BRAF mutation analysis in circulating free tumor DNA of melanoma patients treated with BRAF inhibitors. *Melanoma Res* 25: 486–495.
- Yoon, SH, Choi, YD, Oh, IJ, Kim, KS, Choi, H, Chang, J et al. (2015). Peptide nucleic acid clamping versus direct sequencing for the detection of EGFR gene mutation in patients with non-small cell lung cancer. *Cancer Res Treat* 47: 661–669.
- Han, HS, Lim, SN, An, JY, Lee, KM, Choe, KH, Lee, KH et al. (2012). Detection of EGFR mutation status in lung adenocarcinoma specimens with different proportions of tumor cells using two methods of differential sensitivity. J Thorac Oncol 7: 355–364.
- Jeong, D, Jeong, Y, Lee, S, Lee, H, Lee, W, Kim, H *et al.* (2012). Detection of BRAF(V600E) mutations in papillary thyroid carcinomas by peptide nucleic acid clamp real-time pcr: a comparison with direct sequencing. *Korean J Pathol* 46: 61–67.
- Nagai, Y. Miyazawa, H, Huqun, Tanaka, T, Udagawa, K, Kato, M et al. (2005). Genetic heterogeneity of the epidermal growth factor receptor in non-small cell lung cancer cell lines revealed by a rapid and sensitive detection system, the peptide nucleic acid-locked nucleic acid PCR clamp. *Cancer Res* 65: 7276–7282.
- Lee, B, Lee, B, Han, G, Kwon, MJ, Han, J and Choi, YL (2014). KRAS mutation detection in non-small cell lung cancer using a peptide nucleic acid-mediated polymerase chain reaction clamping method and comparative validation with next-generation sequencing. *Korean J Pathol* 48: 100–107.
- Narumi, S, Matsuo, K, Ishii, T, Tanahashi, Y and Hasegawa, T (2013). Quantitative and sensitive detection of GNAS mutations causing mccune-albright syndrome with next generation sequencing. *PLoS One* 8: e60525.
- Shi, H, Yang, F, Li, W, Zhao, W, Nie, K, Dong, B et al. (2015). A review: fabrications, detections and applications of peptide nucleic acids (PNAs) microarray. *Biosens Bioelectron* 66: 481–489.

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

© CS Hong et al. (2016)