Universal and High-fidelity DNA Single Nucleotide Polymorphism Detection based on CRISPR/Cas12a Biochip

Yong Chen [a], Yixin Mei [a], Xingyu Jiang [ab]*

[a] Department of Biomedical Engineering, Southern University of Science and Technology, No 1088, Xueyuan Rd., Xili, Nanshan District, Shenzhen, Guangdong, 518055, P. R. China
[b] Shenzhen Bay Laboratory, Shenzhen, Guangdong 518055, China

Corresponding Author
E-mail: jiang@sustech.edu.cn

Table of Content

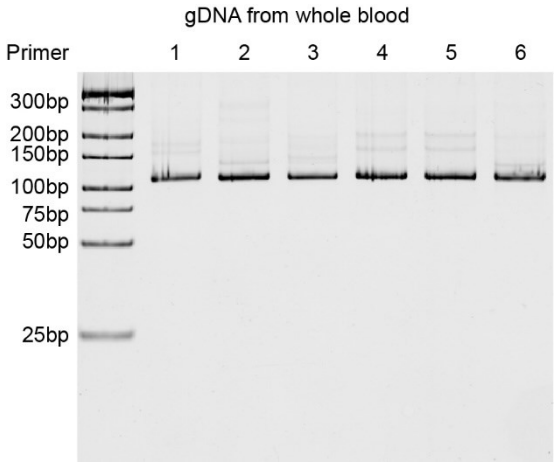


Figure S1. The gel electrophoretic results of the amplified products of the standard DNA using six designed primers. The results show that all six forward primers obtained the pure target band for the synthesized standard samples.

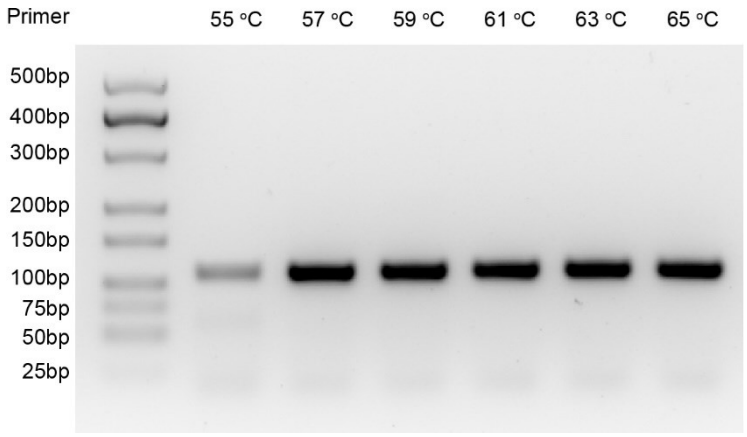


Figure S2. The gel electrophoresis results of the amplification products at different annealing temperatures indicate that the annealing temperature of 59 °C can further improve the amplification efficiency of primer No. 6.

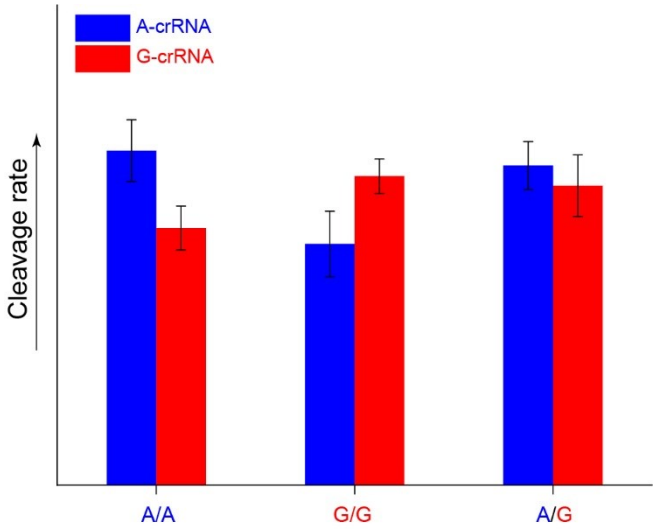


Figure S3. The CRISPR/Cas12a system's detection results for standard samples do not distinguish the SNP genotype. The main reason is that Cas12a nuclease activated by complementary reverse ssDNA has no single nucleotide sensitivity.

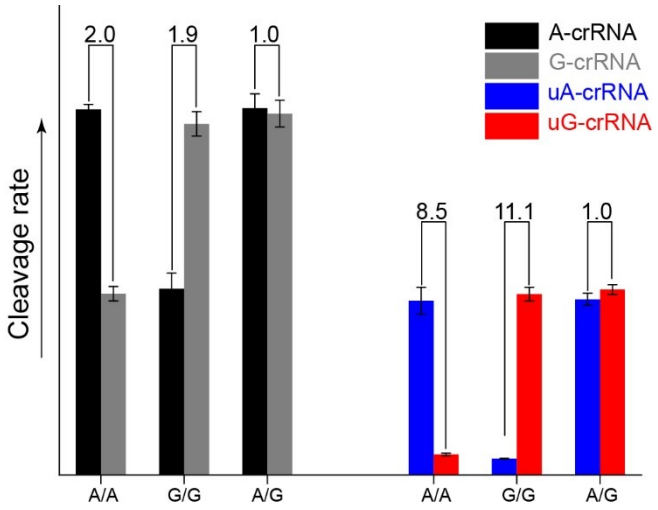


Figure S4. The *trans*-cleavage rate of the CRISPR/Cas12a system under the different number of mismatched nucleotides between crRNA and target DNA. The result showed that the signal difference of the CRISPR/Cas12a system improved from 2-fold to nearly 10-fold when we introduced an addition nucleotide mismatch.

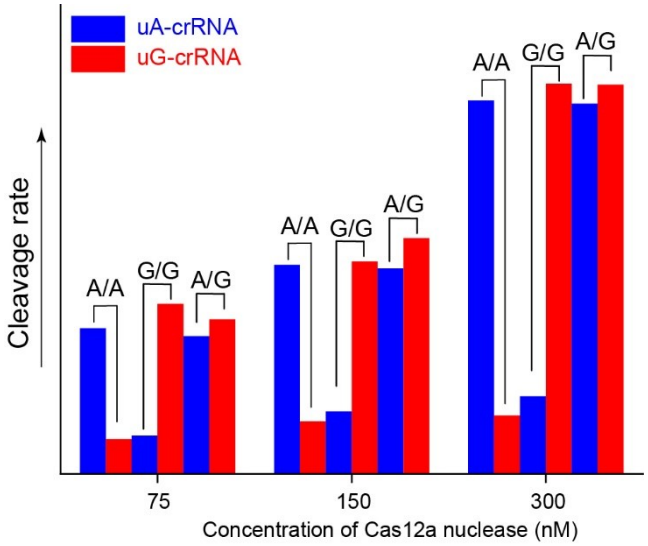


Figure S5. The CRISPR/Cas12a detection result under different concentrations of the Cas12a nuclease. These results indicated that the signal-to-noise ratio (SNR) of the CRISPR/Cas12a system significantly improved with the increasing dosage of the Cas12a nuclease.

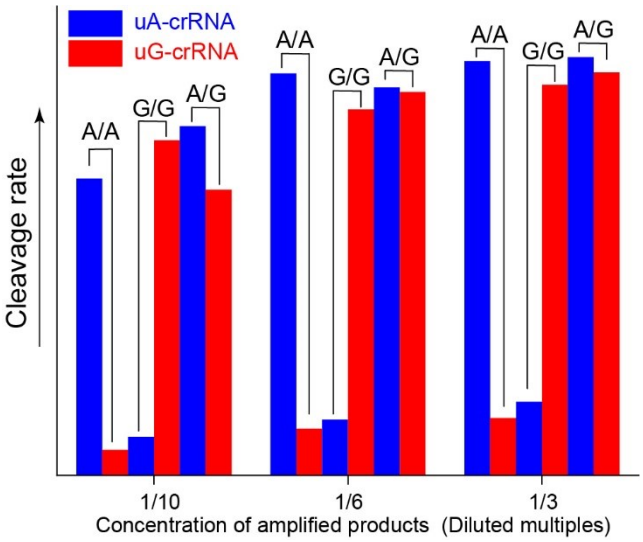


Figure S6. The CRISPR/Cas12a detection result under different concentrations of the amplified product. These results indicated that the detection system is almost unaffected by the concentration of amplified products when the concentration of the Cas12a nuclease is excessive (300 nM > 200 nm, the theoretical maximum concentration of the amplified products in our experiment).

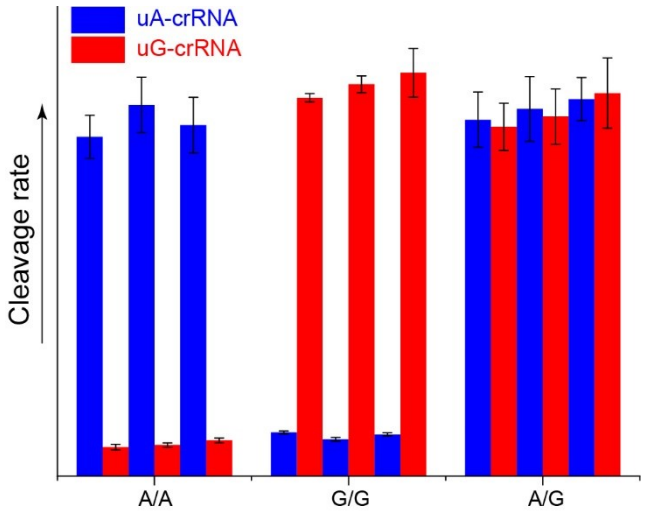


Figure S7. The three repeated genotyping results on the CRISPR/Cas12a biochip. These results indicated that the CRISPR/Cas12a biochip has good stability and reproducibility.

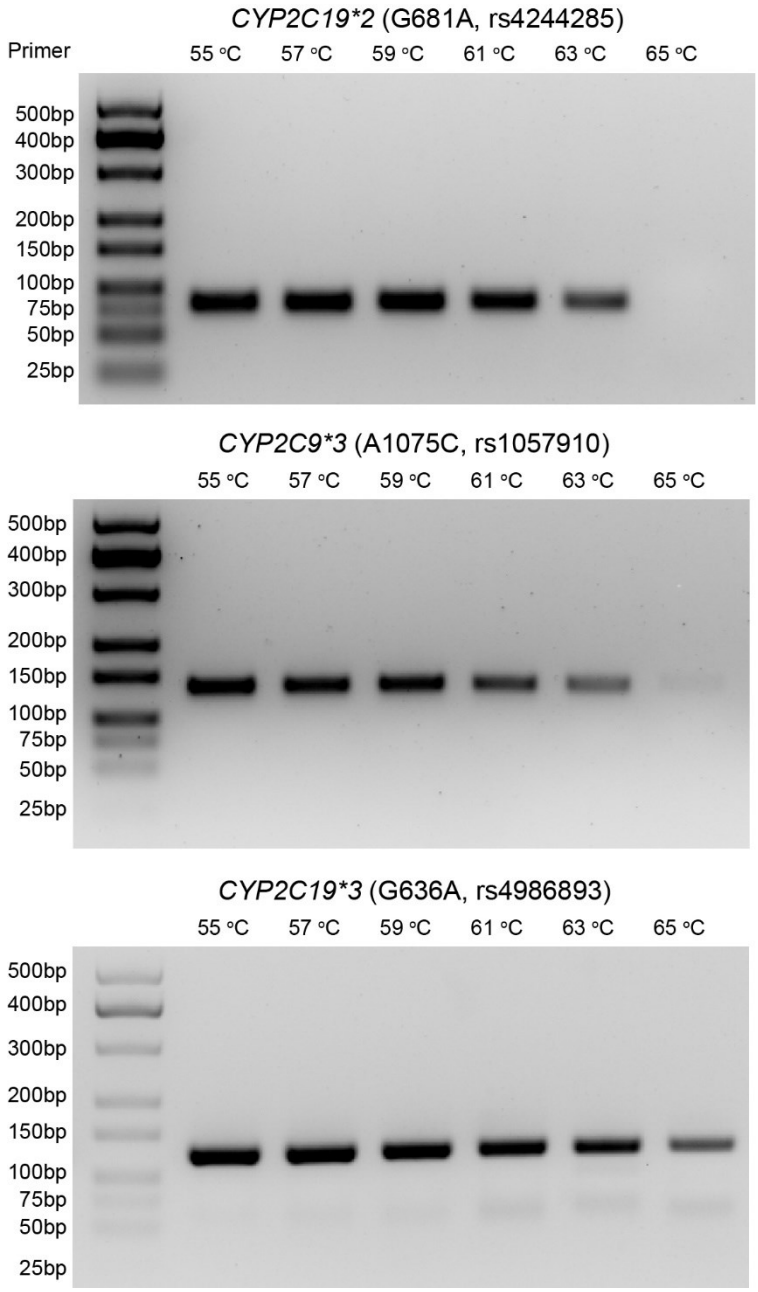


Figure S8. The gel electrophoresis results of the amplification products at different annealing temperatures for *CYP2C19*2* (G681A, rs4244285), *CYP2C9*3* (A1075C, rs1057910), *CYP2C19*3* (G636A, rs4986893) gene, respectively.

Table S1. SNP detection system for *CYP1A1* (A4889G, rs1048943) gene

rs1048943 A/A sequence	CACCCCTGATGGTGCTATCGACAAGGTGTTAAGTGAGAAGGTGATTATCTTTG GCATGGGCAAGCGGAAGTGTATCGGTGAGACCATTGCCCGCTGGGAGGTCTT TCTCTT
rs1048943 G/G sequence	CACCCCTGATGGTGCTATCGACAAGGTGTTAAGTGAGAAGGTGATTATCTTTG GCATGGGCAAGCGGAAGTGTATCGGTGAGACCGTTGCCCGCTGGGAGGTCTT TCTCTT
Forward Primer	AAGAGAAAGACCTCCCAGCGGGCAA
Forward Primer 1	AAGAGAAAGACCTCCCAGCGGTTTA
Forward Primer 2	AAGAGAAAGACCTCCCAGCGTTTAA
Forward Primer 3	AAGAGAAAGACCTCCCAGCTTTCAA
Forward Primer 4	AAGAGAAAGACCTCCCAGTTTGCAA
Forward Primer 5	AAGAGAAAGACCTCCCATTTGGCAA
Forward Primer 6	AAGAGAAAGACCTCCCTTTGGGCAA
Reverse Primer	CACCCCTGATGGTGCTATCGACAAG
Forward Primer-L	CTACCTGAACGGTTTCTCAC
Reverse Primer-L	CTGCATTTGGAAGTGCTC
A-crRNA	UAAUUUCUACUAAGUGUAGAUCAAUGGUCUCACCGAUACAC
G-crRNA	UAAUUUCUACUAAGUGUAGAUCAACGGUCUCACCGAUACAC
uA-crRNA	UAAUUUCUACUAAGUGUAGAUCAuUUGUCUCACCGAUACAC
uG-crRNA	UAAUUUCUACUAAGUGUAGAUCAuCUGUCUCACCGAUACAC
Tag ssDNA	FAM-TTTATT-BHQ1

Primers containing the PAM sequences are defined as forwarding primers.

The mutation site on the target DNA is marked by red.

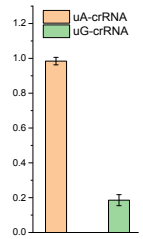
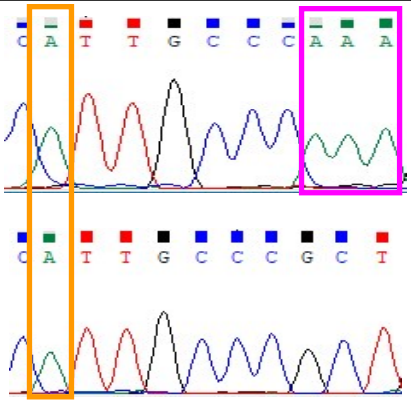
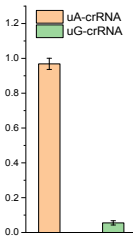
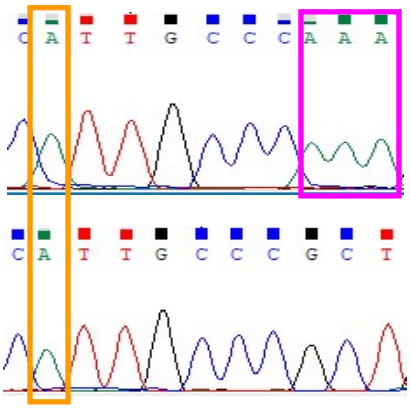
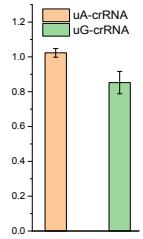
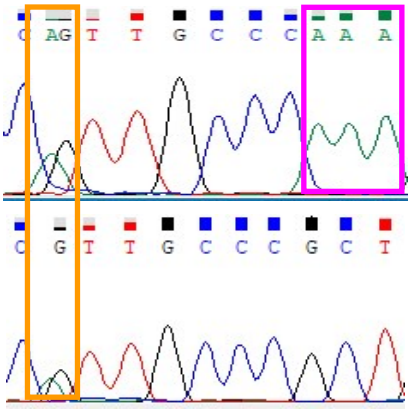
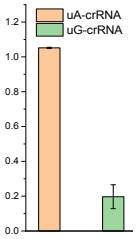
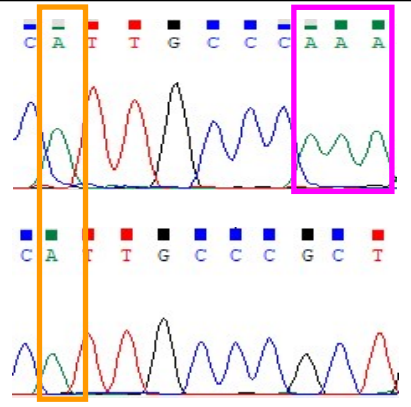
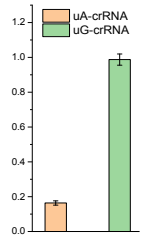
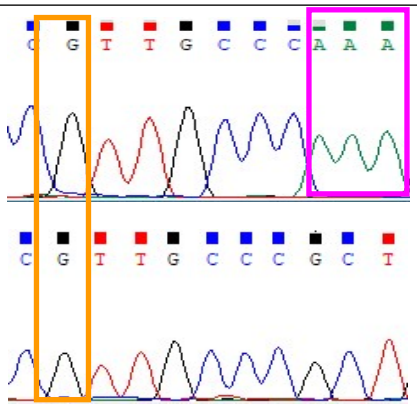
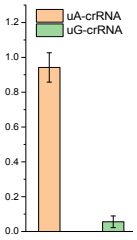
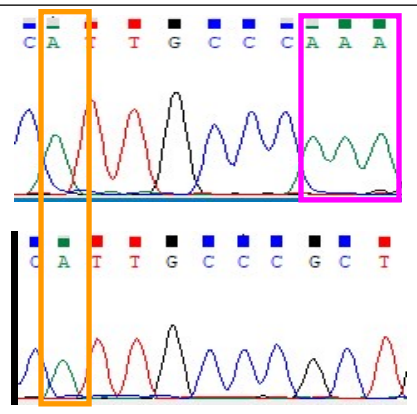
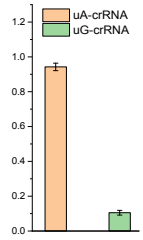
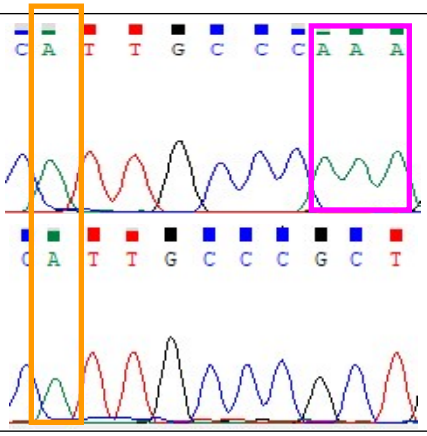
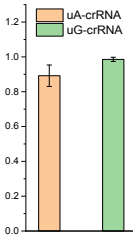
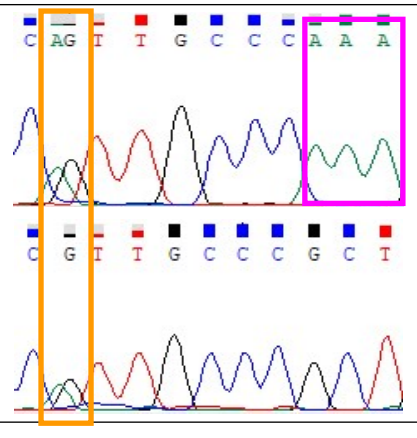
The T bases introduced for PAM sequences are marked by green.

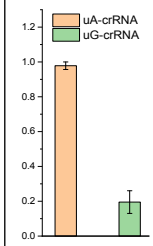
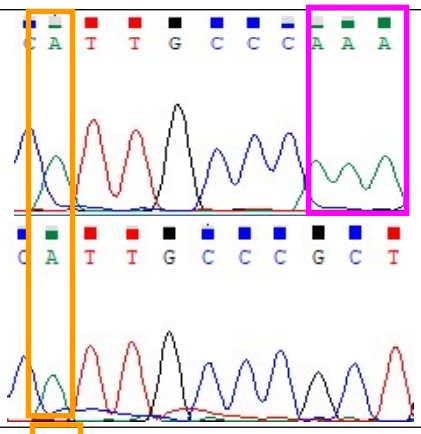
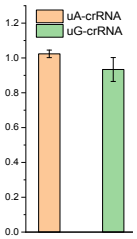
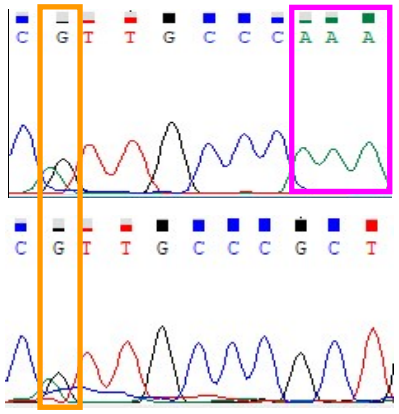
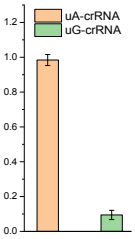
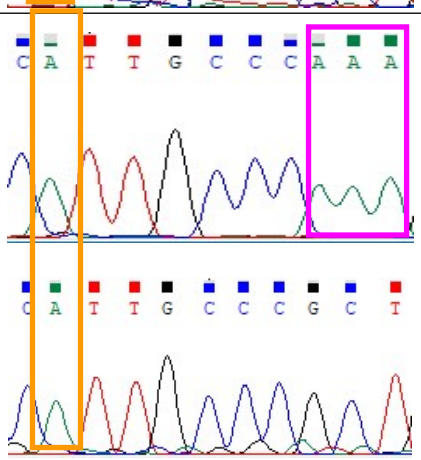
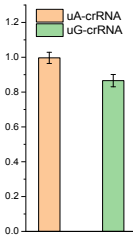
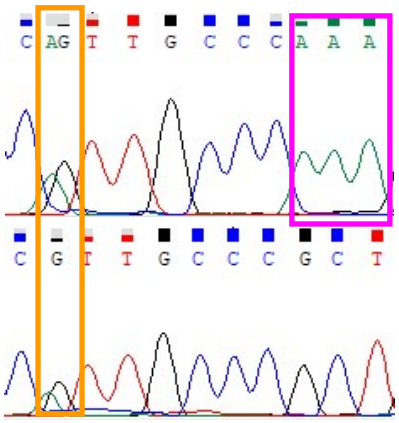
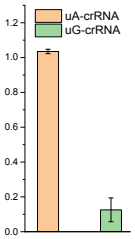
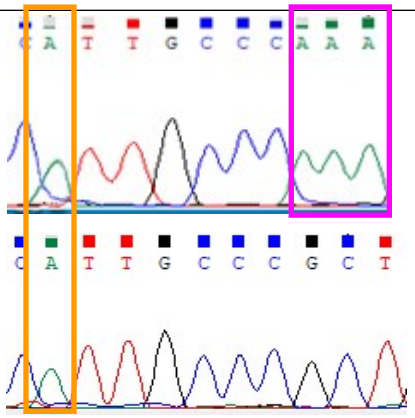
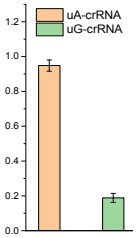
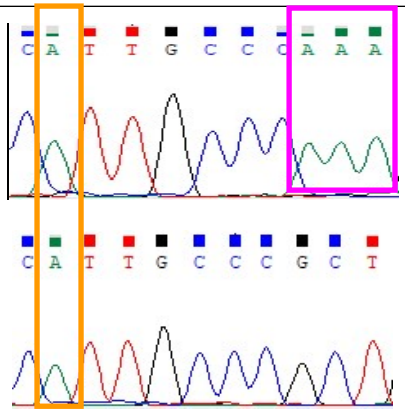
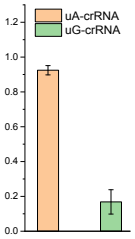
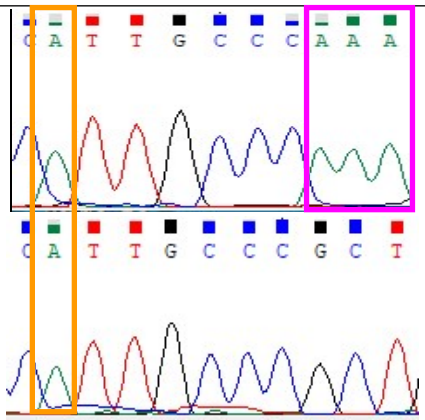
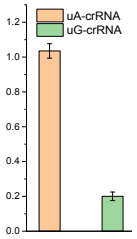
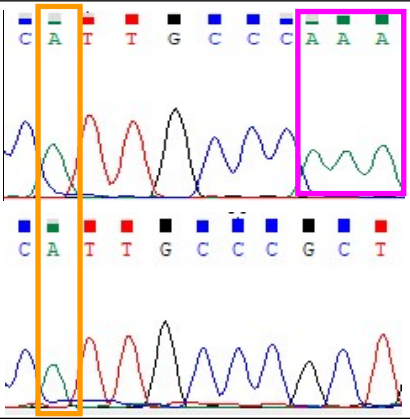
The mutation sites on crRNA are marked by blue.

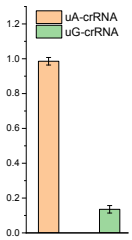
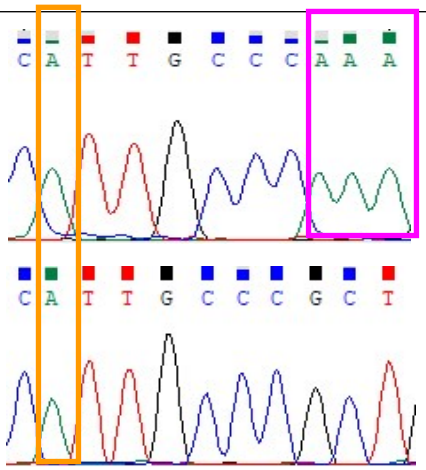
The additional mismatches base on crRNA is marked orange "u".

"Forward Primer-L" and "Reverse Primer-L" were used to amplify 246 bp of long amplification products for reverse pyrosequencing.

Table S2. The result of CRISPR/Cas12a biochip and corresponding pyrosequencing of seventeen human blood samples for the *CYP1A1* (A4889G, rs1048943) gene.

No.	CRISPR	Pyrosequencing Up: forward/reverse primers-PAM Down: forward/reverse primers-L	No.	CRISPR	Pyrosequencing Up: forward/reverse primers-PAM Down: forward/reverse primers-L
1 A/A			2 A/A		
3 A/G			4 A/A		
5 G/G			6 A/A		
7 A/A			8 A/G		

9 A/A			10 A/G		
11 A/A			12 A/G		
13 A/A			14 A/A		
15 A/A			16 A/A		

17 A/A					

The orange curve frame indicates the type of nucleotide at the mutation site.

The purple curve frame indicates the inserted PAM sequences.

The "forward/reverse primer-L" and "forward/reverse primer-PAM" were used for amplification and pyrosequencing. The result indicated that the insertion of PAM sequences does not interfere with the genotypes.

Table S3. SNP detection system for *CYP2C19*2* (G681A, rs4244285) gene

rs4244285 G/G sequence	TTCCCACTATCATTGATTATTTCCCGGGAACCCATAACAAATTACTTAAAAAC CTTGCTTTTATGGAAAGTGATATTTTGG
rs4244285 A/A sequence	TTCCCACTATCATTGATTATTTCCCAGGAACCCATAACAAATTACTTAAAAAC CTTGCTTTTATGGAAAGTGATATTTTGG
Forward Primer	TTCCCACTATCATTGATTATTTCCC
Reverse Primer	CCAAAATATCACTTTCCATAAAAGCAAG
Forward Primer-L	AATTACAACCAGAGCTTGGC
Reverse Primer-L	ATATCACTTTCCATAAAAGCAAG
uG-crRNA	UAAUUUCUACUAAGUGUAGAUCuGGGAACCCAUAACAAAUU
uA-crRNA	UAAUUUCUACUAAGUGUAGAUCuAGGAACCCAUAACAAAUU
Tag ssDNA	FAM-TTTATT-BHQ1

Primers containing the PAM sequences are defined as forwarding primers.

The mutation site on the target DNA is marked by red.

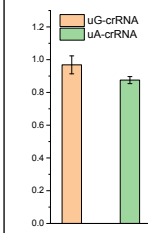
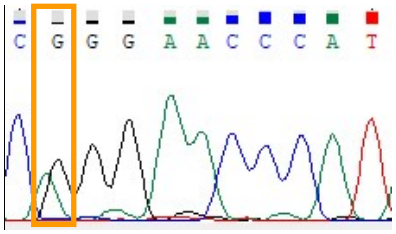
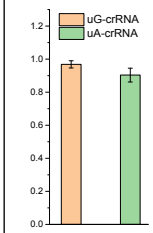
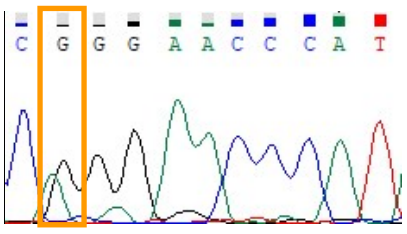
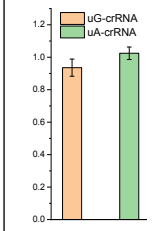
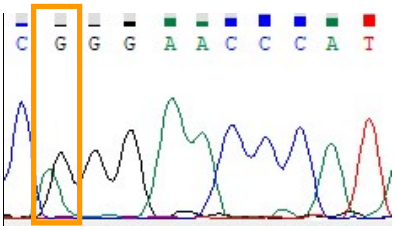
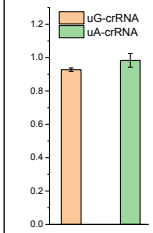
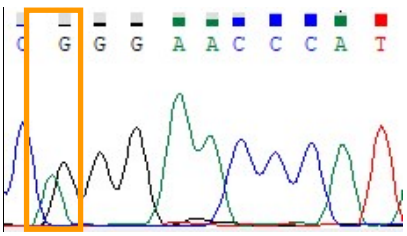
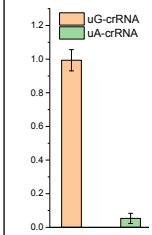
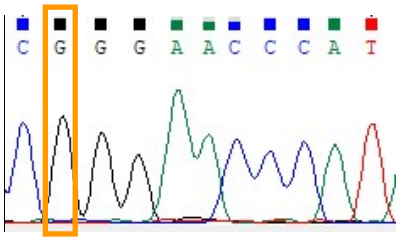
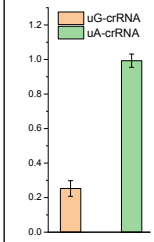
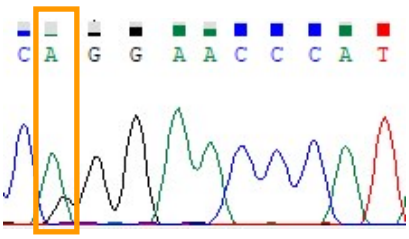
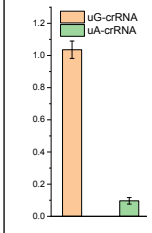
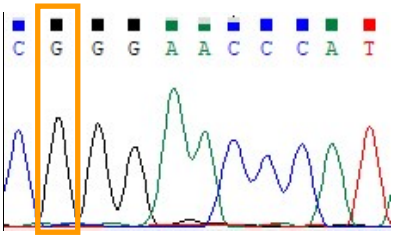
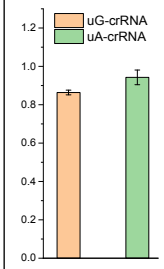
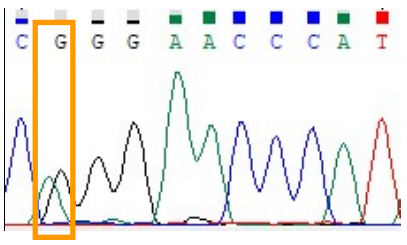
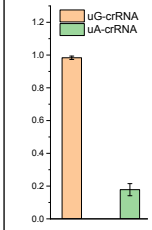
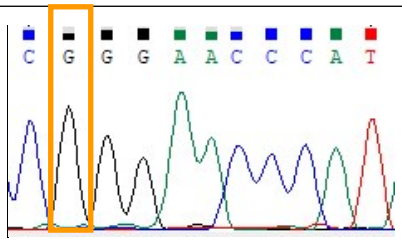
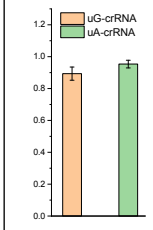
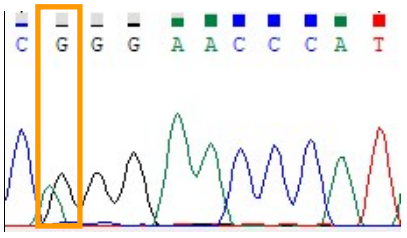
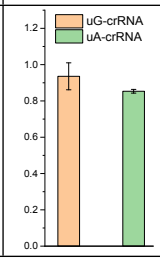
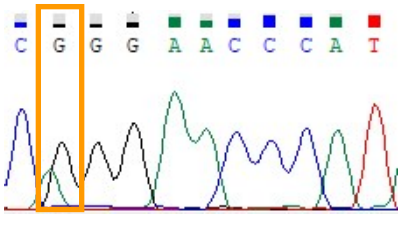
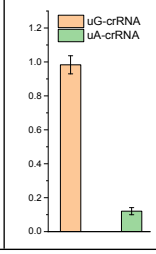
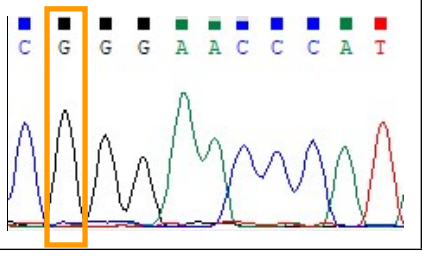
The T bases introduced for PAM sequences are marked by green.

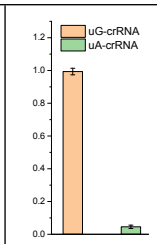
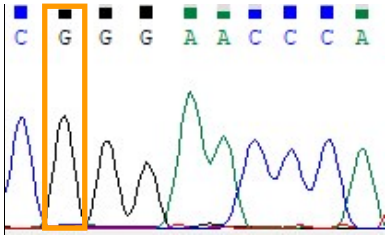
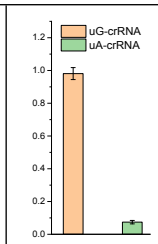
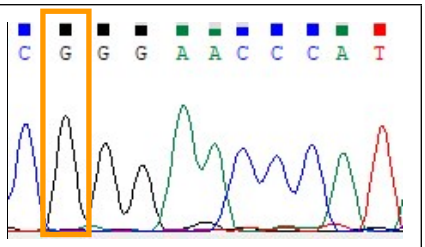
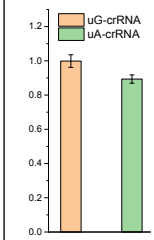
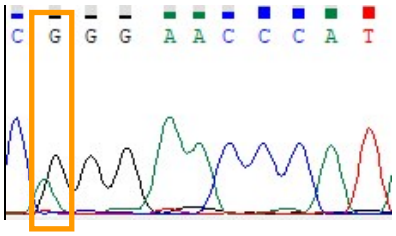
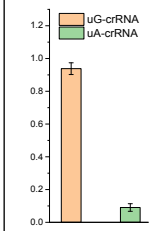
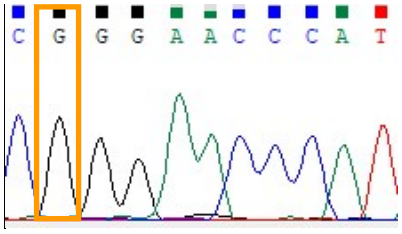
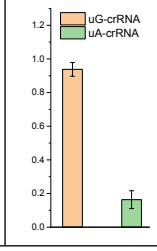
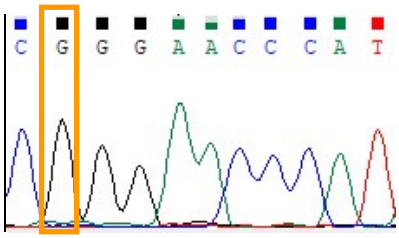
The mutation sites on crRNA are marked by blue.

The additional mismatches base on crRNA is marked orange "u".

"Forward Primer-L" and "Reverse Primer-L" were used to amplify 168 bp of long amplification products for forward pyrosequencing.

Table S4. The result of CRISPR/Cas12a biochip and corresponding pyrosequencing of seventeen human blood samples for the *CYP2C19*2* (G681A, rs4244285) gene.

No.	CRISPR	pyrosequencing	No.	CRISPR	pyrosequencing
1 G/A			2 G/A		
3 G/A			4 G/A		
5 G/G			6 A/A		
7 G/G			8 G/A		
9 G/G			10 G/A		
11 G/A			12 G/G		

13 G/G			14 G/G		
15 G/A			16 G/G		
17 G/G					

The orange curve frame indicates the type of nucleotide at the mutation site.

For pyrosequencing, the "Forward Primer-L" and "Reverse Primer-L" were used for amplification to obtain long sequences, so the pyrosequencing results did not contain PAM sequences.

Table S5. SNP detection system for *CYP2C9*3* (A1075C, rs1057910) gene

rs1057910 A/A sequence	CCAGGAAGAGATTGAACGTGTGATTGGCAGAAACCGGAGCCCCTGCATGCAA GACAGGAGCCACATGCCCTACACAGATGCTGTGGTGCACGAGGTCCAGAGAT ACATTGACCTTCTCCCCACCAGCCTGCC
rs1057910 C/C sequence	CCAGGAAGAGATTGAACGTGTGATTGGCAGAAACCGGAGCCCCTGCATGCAA GACAGGAGCCACATGCCCTACACAGATGCTGTGGTGCACGAGGTCCAGAGAT ACCTTGACCTTCTCCCCACCAGCCTGCC
Forward Primer	GGCAGGCTGGTGGGGATTTGGTCAA
Reverse Primer	CCAGGAAGAGATTGAACGTGTGATTG
uA-crRNA	UAAUUUCUACUAAGUGUAGAUGUCAuUGUAUCUCUGGACCU
uC-crRNA	UAAUUUCUACUAAGUGUAGAUGUCAuGGUAUCUCUGGACCU
Tag ssDNA	FAM-TTTATT-BHQ1

Primers containing the PAM sequences are defined as forwarding primers.

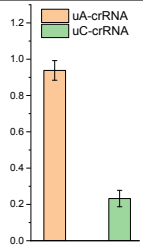
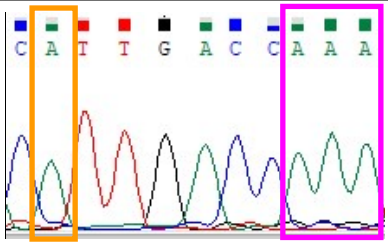
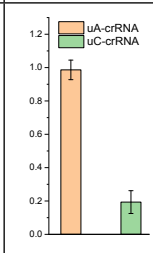
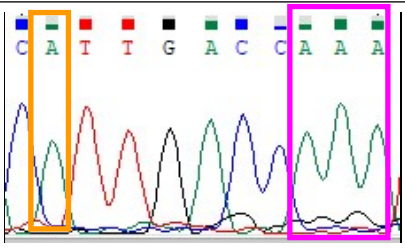
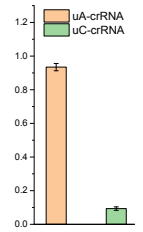
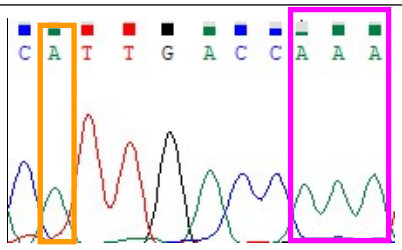
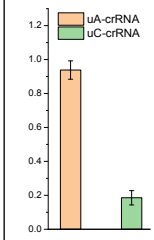
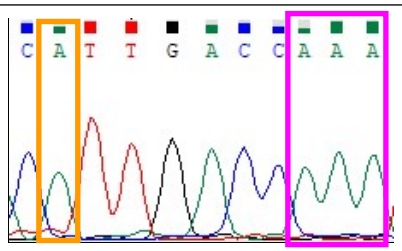
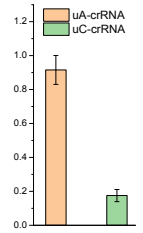
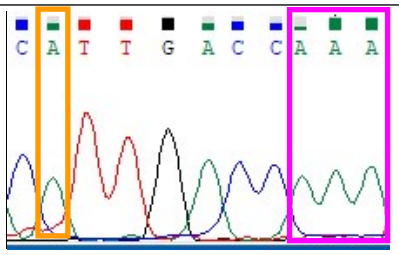
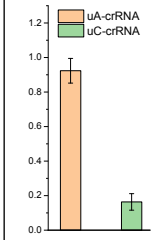
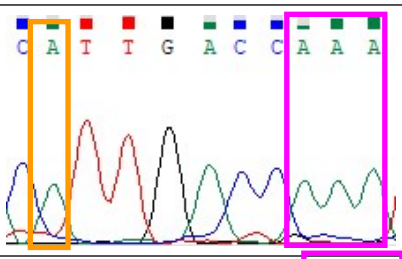
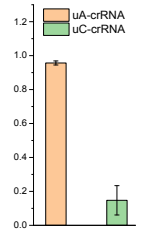
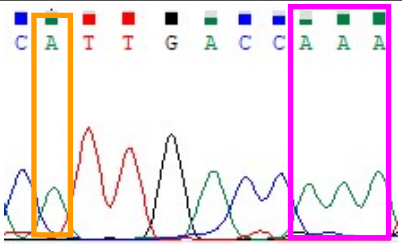
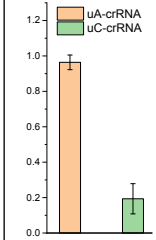
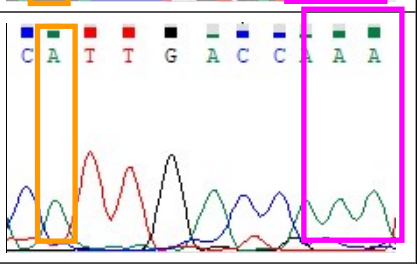
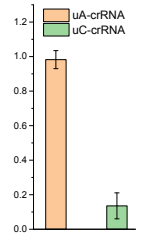
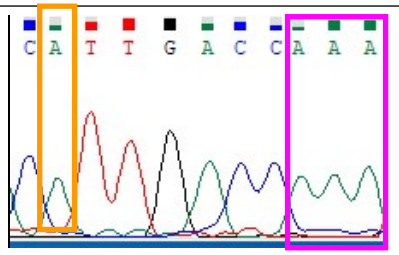
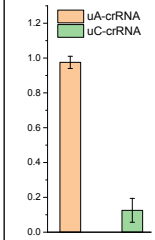
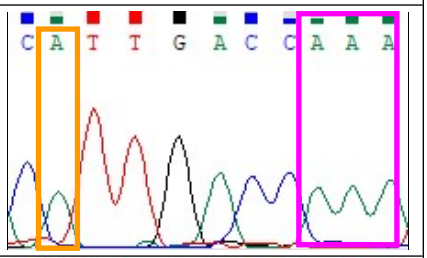
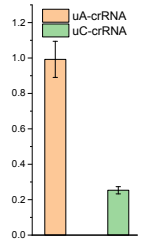
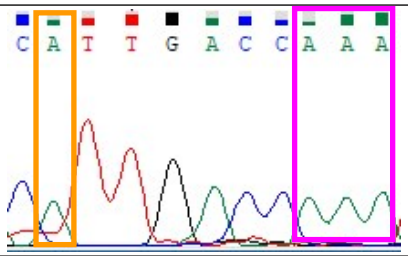
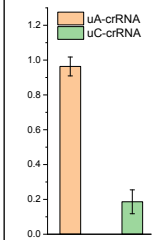
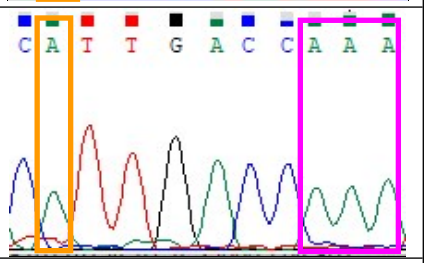
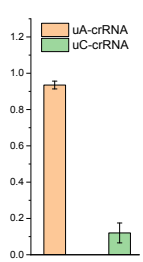
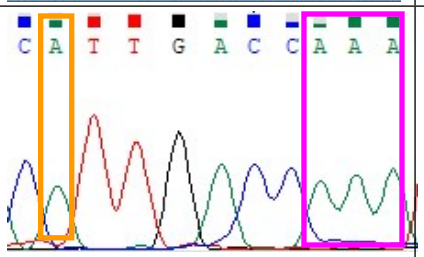
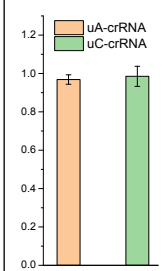
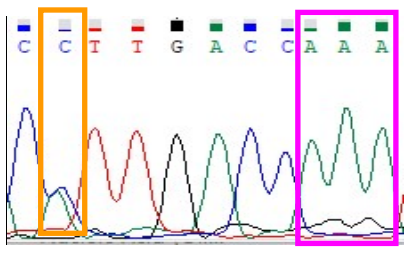
The mutation site on the target DNA is marked by red.

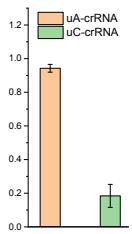
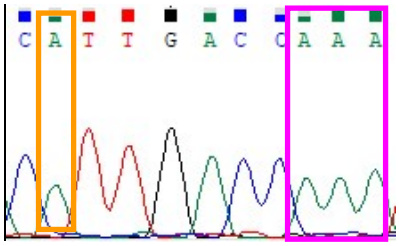
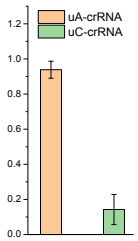
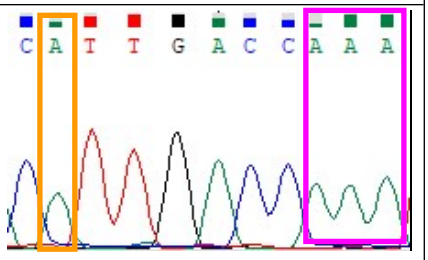
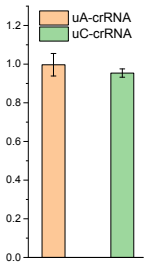
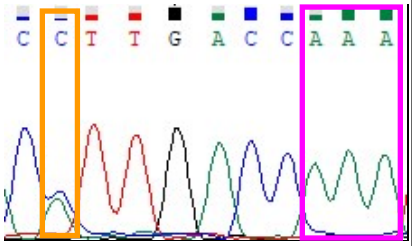
The T bases introduced for PAM sequences are marked by green.

The mutation sites on crRNA are marked by blue.

The additional mismatches base on crRNA is marked orange "u".

Table S6. The result of CRISPR/Cas12a biochip and corresponding pyrosequencing of seventeen human blood samples for the *CYP2C9*3* (A1075C, rs1057910) gene.

No.	CRISPR	pyrosequencing	No.	CRISPR	pyrosequencing
1 A/A			2 A/A		
3 G/G			4 G/G		
5 G/G			6 G/G		
7 G/G			8 G/G		
9 G/G			10 G/G		
11 G/A			12 G/A		
13 G/G			14 A/C		

15 G/G			16 G/G		
17 G/G					

The orange curve frame indicates the type of nucleotide at the mutation site.

The purple curve frame indicates the inserted PAM sequences.

Table S7. SNP detection system for *CYP2C19*3* (G636A, rs4986893) gene

rs4986893 G/G sequence	ACATCAGGATTGTAAGCACCCCCTGGATCCAGGTAAGGCCAAGTTTTTTGCTT CCTGAGAAACCACTTACAGTCTTTTTTTCTGGGAAATCCAAAATTCTATATTG ACCAAGCCCTGAAGTACAT
rs4986893 A/A sequence	ACATCAGGATTGTAAGCACCCCCTGAATCCAGGTAAGGCCAAGTTTTTTGCTT CCTGAGAAACCACTTACAGTCTTTTTTTCTGGGAAATCCAAAATTCTATATTG ACCAAGCCCTGAAGTACAT
Forward Primer	ACATCAGGATTGTAAGTTTCCCCTG
Reverse Primer	ATGTACTTCAGGGCTTGGTCAATA
aG-crRNA	UAAUUUCUACUAAGUGUAGAUCCCUuGAUCCAGGUAAGGCC
aA-crRNA	UAAUUUCUACUAAGUGUAGAUCCCUuAAUCCAGGUAAGGCC
Tag ssDNA	FAM-TTTATT-BHQ1

Primers containing the PAM sequences are defined as forwarding primers.

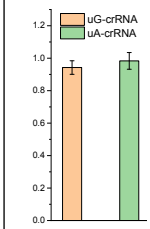
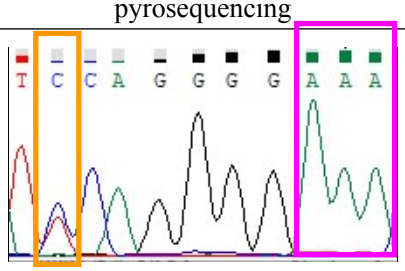
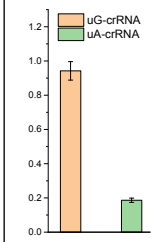
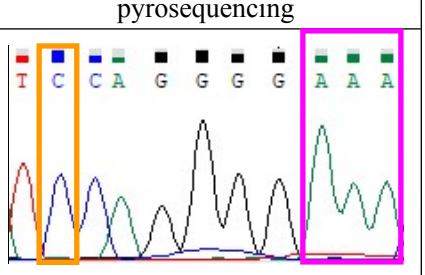
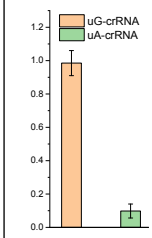
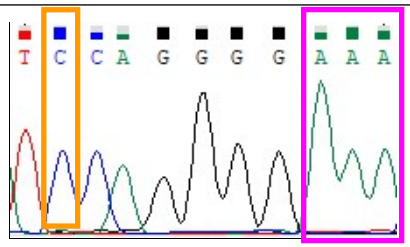
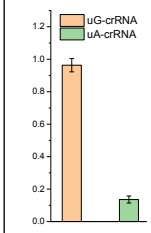
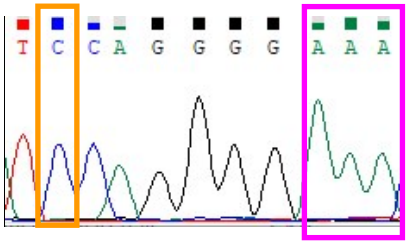
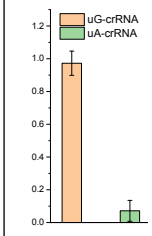
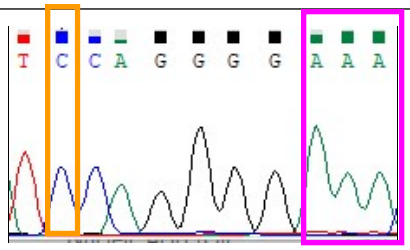
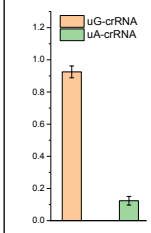
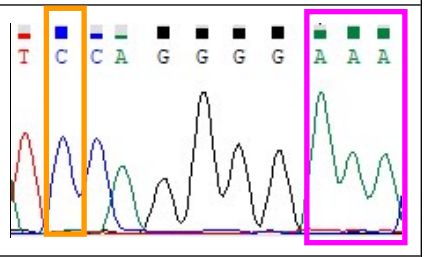
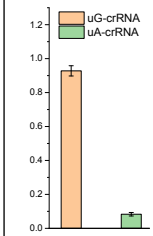
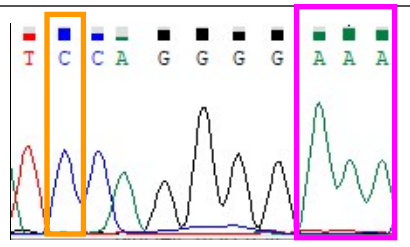
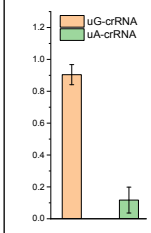
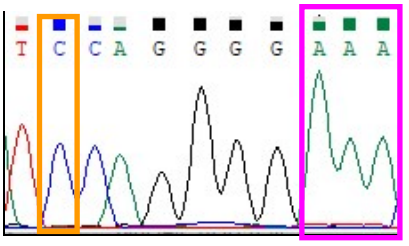
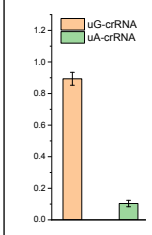
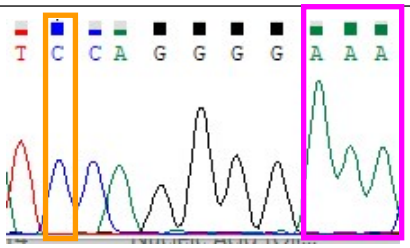
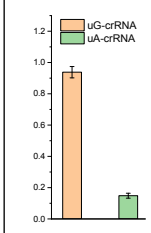
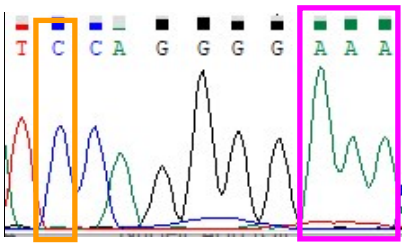
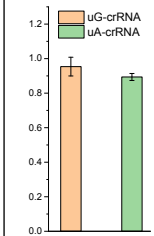
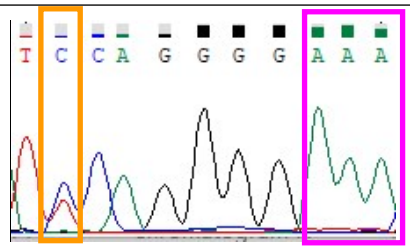
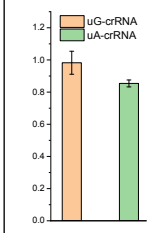
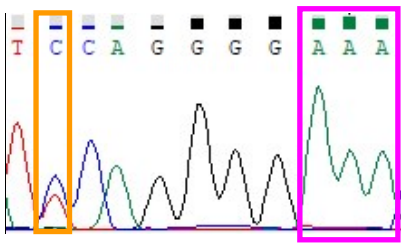
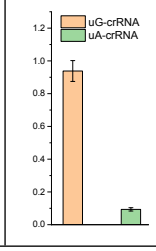
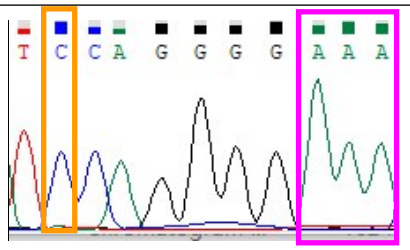
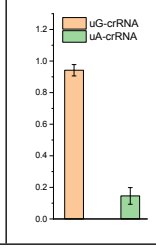
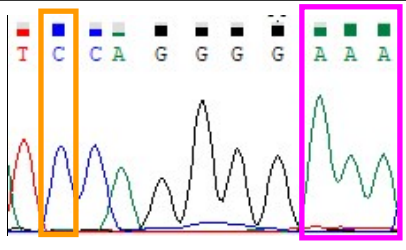
The mutation site on the target DNA is marked by red.

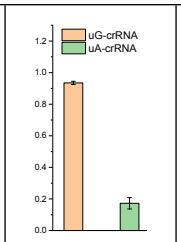
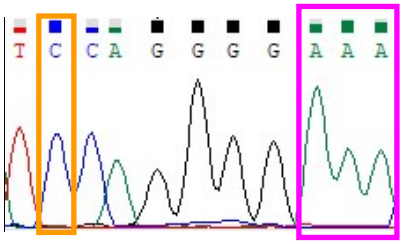
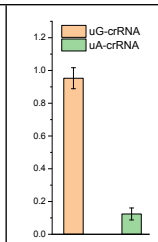
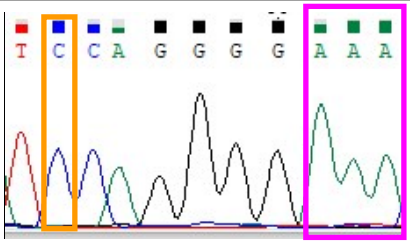
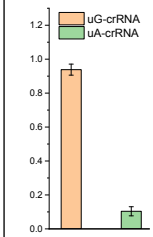
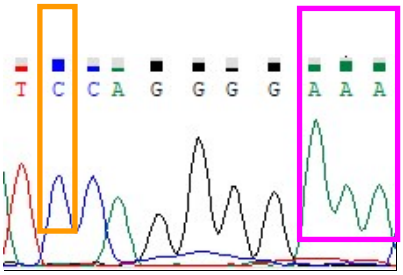
The T bases introduced for PAM sequences are marked by green.

The mutation sites on crRNA are marked by blue.

The additional mismatches base on crRNA is marked orange "u".

Table S8. The result of CRISPR/Cas12a biochip and corresponding pyrosequencing of seventeen human blood samples for the *CYP2C19*3* (G636A, rs4986893) gene.

No.	CRISPR	pyrosequencing	No.	CRISPR	pyrosequencing
1 G/A			2 G/G		
3 G/G			4 G/G		
5 G/G			6 G/G		
7 G/G			8 G/G		
9 G/G			10 G/G		
11 G/A			12 G/A		
13 G/G			14 G/G		

15 G/G			16 G/G		
17 G/G					

The orange curve frame indicates the type of nucleotide at the mutation site.

The purple curve frame indicates the inserted PAM sequences.