Transcription of AAT•ATT Triplet Repeats in *Escherichia coli* Is Silenced by H-NS and IS1E Transposition

Xuefeng Pan¹,²,³*, Yuanhong Liao¹,³*, Yunmeng Liu¹,³, Peng Chang¹,³, Lingni Liao¹, Li Yang¹, Hongquan Li³

¹ School of Life Science, Beijing Institute of Technology, Beijing, China, ² Institute of Cell and Molecular Biology, the University of Edinburgh, Edinburgh, United Kingdom, ³ Health Science Centre, Hebei University, Baoding, China

Abstract

**Background:** The trinucleotide repeats AAT•ATT are simple DNA sequences that potentially form different types of non-B DNA secondary structures and cause genomic instabilities in vivo.

**Methodology and Principal Findings:** The molecular mechanism underlying the maintenance of a 24-triplet AAT•ATT repeat was examined in *E.coli* by cloning the repeats into the EcoRI site in plasmid pUC18 and into the attB site on the *E.coli* genome. Either the AAT or the ATT strand acted as lagging strand template in a replication fork. Propagations of the repeats in either orientation on plasmids did not affect colony morphology when triplet repeat transcription using the lacZ promoter was repressed either by supplementing Lac⁴ in trans or by adding glucose into the medium. In contrast, transparent colonies were formed by inducing transcription of the repeats, suggesting that transcription of AAT•ATT repeats was toxic to cell growth. Meanwhile, significant IS1E transposition events were observed both observed between the repeat regions proximal to the promoter side, the promoter region of the lacZ gene, and into the AAT•ATT region itself. Transposition reversed the transparent colony phenotype back into healthy, convex colonies. In contrast, transcription of an 8-triplet AAT•ATT repeat in either orientation on plasmids did not produce significant changes in cell morphology and did not promote IS1E transposition events. We further found that a role of IS1E transposition into plasmids was to inhibit transcription through the repeats, which was influenced by the presence of the H-NS protein, but not of its parologue StpA.

**Conclusions and Significance:** Our findings thus suggest that the longer AAT•ATT triplet repeats in *E.coli* become vulnerable after transcription. H-NS and its facilitated IS1E transposition can silence long triplet repeats transcription and preserve cell growth and survival.

Introduction

AAT•TTA triplet repeats are types of trinucleotide repeats that are highly repetitive in the human genome [1,2], in contrast, they are much less abundant in many prokaryotic genomes, such as in *E.coli*, showing a biased distribution towards eukaryotic genomes [3,4,5,6,7]. Further, AAT•TTA repeats are also found to be over-represented in intronic and intergenic regions, but underrepresented in exons and UTRs, showing a nonrandom nature of distribution in an individual genome [1,2].

AAT•TTA repeats have been characterized in vitro to form various types of non-B DNA secondary structures, including hairpin, triplex, non-H DNA [8,9,10,11], which potentially challenge the stable maintenance of the repeats in genomes [8].

Consisting with this, it has been found in prokaryotes that certain proteins known as chromosome structural proteins bind preferentially to AT-rich DNA that is normally seen in promoter region, and which may make the DNA segments inactive to avoid forming non-B secondary structure during DNA transcription [12]. H-NS and its parologue StpA are two of the main chromosome structural proteins in *E.coli*, which bind AT-rich DNA sequences with overlapped specificity; H-NS binds to more than 1000 genes when repressing the transcription of the target DNA region [12].

The non-B structures formed by AAT•TTA triplet repeats were found to be similar to those of non-B secondary structures formed by disease causing trinucleotide repeats such as CAG•CTG, CCG•CCG and GAA•TTC, which are associated with more than 40 human genetic diseases, including Huntington’s disease or fragile X syndrome [8,13]. Consistent with this, a typical expansion feature of AAT•TTA triplet repeats was found in an in vitro amplification assay [9,10,13,14,15,16,17,18]. Moreover, in human populations similar instabilities of AAT•TTA triplet repeats were recently found to be associated with the high IgE blood syndrome in Chinese children [19], with schizophrenia [20,21], cocaine addiction [22], and with the high prevalence of depression in adult Parkinson’s disease patients [23]. More
strikingly, a similar triplet repeat expansion was also found to be associated with the propagation of a 33-bp AT-rich repeat, which displayed chromosomal fragilites in humans, such as FRA3B, FRA16B, and FRA10B [24,25]. However, the maintenance and biased distribution of AAT•TTA repeats in different chromosomal regions in a genome or in different genomes are still poorly understood. To gain some understandings of the maintenance of AAT•TTA triplet repeats in vivo and its biased distribution within a genome or in genomes, we performed experiments using a 24-triplet AAT•TTA repeats in the *E. coli* model system. We found that propagation of the repeats on plasmids was overall normal, but that transcription of the repeats in either orientation with either AAT or ATT serving as transcribing template induced significant changes in colony morphology, leading to formation of convex colonies from normal colonies, which then progressively transformed into transparent colonies. Interestingly, we found that IS1E transposition from the chromosome to either the proximal site of the promoter of lacZ in the plasmids or into the AAT•TTA triplet repeats themselves were capable of reversing colony changes by repressing transcription of the repeats. And we further found that the AT-rich repeats binding protein H-NS, a chromosomal structuring protein, but not its paralogoue StpA, played dual roles in triplet repeats binding and in facilitating IS1E transposition into the repeats containing plasmid, which ultimately prevented transcription of the triplet repeats.

Results

Transcription of a 24-triplet AAT•TTA repeat on a plasmid induces morphological changes of bacterial colonies from convex to transparent

Plasmids carrying a 24-triplet AAT•TTA repeat with either the AAT or the ATT strand as lagging strand template for DNA replication were initially constructed as described in the materials and methods. They were named as pAAT24 and pTTA24, respectively. Plasmid pAAT24 was a deletion product occasionally obtained when propagating pAAT24 in a *rec* mutant, and the pTTA24 plasmid was obtained by reversing the AAT\(_\text{24}\) into the opposite orientation.

When the pAAT24 and pTTA24 plasmids were propagated in the *E. coli* JM83 wildtype strain, normal convex colonies on LB plates containing ampicillin were formed, but the colonies started to become flat when plates were left at 4°C overnight (Figure 1). Almost all flat colonies became concave, and then progressively turned into transparent colonies. In contrast, a similar phenomenon was not observed with JM83 cells when they propagated pAAT\(_\text{18}\), pTTA\(_\text{18}\), and pUC18 plasmids growing under the same conditions (data not shown). This indicated that only colonies carrying longer AAT•TTA triplet repeats suffered morphological changes when grown overnight.

Several processes may be responsible for the observed alterations in colony morphology, including triplet repeats transcription from the lacZ promoter (Figure 1A) [26], DNA double strand breaks caused by formation of non-B secondary structures, such as non-H structure, which may be targeted by certain DNA structure specific nucleases [27,28], or activation of the cryptic phage\(\Phi\) 80 in the JM83 chromosome, leading to cell lysis.

To test these possibilities, plasmids pAAT\(_\text{24}\) and pTTA\(_\text{24}\) were also transformed into JM83 mutants defective in homologous recombination, such as Δhns, ΔrecB, *recF::Tn10 Kan\(_r\)*, and ΔhnsΔrecBC. We wanted to test if the colony morphology changes may be hinted at AAT•TTA triplet repeats breaks in plasmids, which requires homologous recombination repair, and which relies on the *rec* gene functions. It was observed that DNA double strand breaks in *E. coli* genome or plasmid affected cell viability and also the colony morphology, leading to plasmid loss or cell death.

However, we found that compared to wild type cells, propagations of pAAT\(_\text{24}\) and pTTA\(_\text{24}\) plasmids in all *rec* mutants tested did not significantly enhance morphological changes (data not shown).

Therefore, we excluded double strand break formation as the dominant cause for the observed changes in colony morphology.

To test whether the genetic background of the *E. coli* host made a difference, the growth and colony morphology of AB1157recF::Tn10 and W3110Δphd::cat cells containing plasmids pAAT\(_\text{24}\) and pTTA\(_\text{24}\) were compared to that of JM83. Although their genetic backgrounds were significantly different from JM83 and no similar cryptic phages were found, AB1157recF::Tn10 and W3110Δphd::cat colonies suffered the same morphological changes as JM83 cells (Figure 1B). These suggested that the observed morphological changes as associated with AAT•TTA triplet repeats only depended on the presence of the plasmids, but not on the genetic background or cryptic phage of the host. Therefore, all subsequent work was done using the JM83 wild type and its derivative strains.

As shown in Figure 1A, AAT•ATT triplet repeats were subcloned at the *EcoRI* site of pUC18, which is located downstream from the lacZ promoter (Figure 1A). Therefore, we wondered whether transcription of the AAT•ATT triplet repeats from the lacZ promoter occurred in the stationary phase, thus causing directly or indirectly colony morphology changes.

To test whether transcription was responsible for the observed changes in colony morphology, we constructed JM83Δhns and JM83ΔstpA::cat mutants defective in chromosomal structuring proteins H-NS and StpA, respectively. We therefore correlated the possible effects of transcription of AAT•ATT triplet repeats in pAAT\(_\text{24}\) and pTTA\(_\text{24}\) with morphological changes in the two strains by supplementing with IPTG, or by co-expressing LacI\(_Q\) from a plasmid. Interestingly, compared to JM83 and its *rec* derivatives, propagation of pAAT\(_\text{24}\) and pTTA\(_\text{24}\) in the JM83Δhns mutant did not provoke changes in colony morphologies, but propagation in JM83ΔstpA::cat did, indicating that H-NS, but not StpA, was required for this process (Figure 1B).

H-NS binds to AT-rich DNA, which may have an effect on AAT•TTA triplet repeats transcription and somehow stimulate morphological changes of the colonies. To test additional effects, pAAT\(_\text{24}\) and pTTA\(_\text{24}\) were transformed into the ΔhnsrecBC double mutant. No significant differences in morphological changes between the single *Δhns* and the ΔhnsrecBC double mutant were observed (data not shown). Therefore, transcription of AAT•TTA triplet repeats did not cause double strand breaks in the presence or absence of the H-NS protein.

Transcription of AAT•ATT triplet repeats stimulates transposition of IS1E from the chromosome to plasmids

During prolonged incubation of morphologically changed colonies, secondary colonies can be slowly regenerated from the transparent cells containing pAAT\(_\text{24}\) and pTTA\(_\text{24}\) plasmids that eventually form healthy colonies in situ (Figure 1). The plasmid DNA was isolated from these regenerated colonies, and analyzed for variations on an agarose gel (Figure 2A). Plasmids recovered from the regenerated colonies were bigger in size, running between the monomeric and the dimeric plasmids (Figure 2A).

Restriction enzyme digestions showed that extra *Pst*I cutting sites were the result of an IS1E insertion into the plasmid.
A. (AAT)$_n$ as template for lagging strand synthesis

B. (ATT)$_n$ as template for lagging strand synthesis

(1)  (2)  (3)  (4)

(5)  (6)  (7)  (8)

(9)  (10)  (11)  (12)

(13)  (14)  (15)  (16)
The IS1E insertions were often on the 5’ side of the flanking sequence of the AAT•ATT triplet repeats, that is, within the promoter region of the lacZ gene (Figure 2B). For example, IS1E insertions occurred with a higher frequency into the “GGAATTGTG” site, where the first adenine base was the starting site for LacZ mRNA transcription. Another high frequency site was “GATTACGAA”, containing the “GAA” of the EcoRI restriction site 5’ to the repeats (Figure 2B). These findings suggested that transcription of the AAT•ATT triplet repeats was responsible for the negative effects on cell growth.

Figure 2. Characterization of the IS1E insertion positions. A) a. Transposition of IS1E elements into plasmids, plasmid isolated from the healthy colonies, and b. Restriction digestion of the plasmid DNA with transposons. Lanes 1, molecular weight (pBR322/BstN1); Lane 2 and 3, pAAT24 and pATT24 without IS1E transposition, digested by PstI; Lane 4, pAAT24, with IS1E transposition, digested by PstI; Lane 5, pAAT24 with IS1E transposition, digested by EcoRI; Lane 6 and 7, pAAT24 with IS1E transposition, digested by PstI and EcoRI respectively; Lane 8 and 9 pAAT24 with IS1E transposition, digested by using PvuII. B) Mapping the IS1E inserting elements in the promoter and the AAT repeats region by DNA sequencing, the positions of the mostly recognized sites were shown.
AAT•ATT triplet repeats serve as targets for IS1E transposition in repeats transcription

Besides recovering IS1E insertions upstream of the AAT•ATT repeats, we also isolated plasmids containing IS1E insertions in the AAT•ATT repeating array itself (Figure 2B and Figure 3). EcoRI fragments of ~1 kb in size containing at least one IS1E element within the AAT•ATT repeats were recovered from both pAAT24 and pTTA24 plasmids (Figure 3A). These findings are of particular interest, as this is the first demonstration of AAT•ATT repeats acting as potential hotspots for transposition in vivo. In this case, we have further found that the insertion of IS1E into AAT•ATT repeats rely on the repeats orientations in transcription. IS1E appeared to favour inserting itself into the AAT•ATT repeats with its unique PstI restriction site distal to the PstI restriction site in the vector when AAT orientation is transcribed, generating larger DNA fragment after PstI digestion (Figure 3A lane 1 and 3, and 3B), while insertion of IS1E in the repeats array will use the opposite direction if the ATT orientation is transcribed, generating small DNA fragment by PstI digestion (Figure 3A lane 5 and 7, and 3B). These suggested that ISIE transposition into AAT•ATT repeats occurred in repeats transcription. Transcription of AAT repeats and ATT repeats may produce distinct repeats DNA conformations that may differently be recognized by H-NS, and therefore differently affect the IS1E transpositions. Based upon the fact that IS1E transposition into AAT•ATT repeats expands the repeats by 3 triplets of either AAT or ATT repeats (9 nts) at both the 5’ and 3’ ends flanking the IS1E insertion, which may imply a potential way of making repeats expansion by certain transposons in trinucleotide repeats [13,29,30,31,32].

Transcription of AAT•ATT triplet repeats is inactivated by IS1E transposition

To understand the biological significance of the IS1E transposition during the transcription of AAT•ATT triplet repeats, we have further analyzed the effect of IS1E transposition on lacZ transcription in the plasmids. To this end, we eliminated the AAT•ATT triplet repeats in pAAT24IS1E-A using EcoRI digestion (as marked in Figure 2B), and subsequently re-ligated into the repeats-free pAAT24IS1E-A vector at the EcoRI restriction site. As determined by DNA sequencing, insertion of an IS1E into pAAT24IS1E-A did not alter the lacZ gene promoter significantly (Figure 2), however, the IS1E element could potentially use its promoter to transcribe both the transposase gene (encoded by it) and the AAT•ATT triplet repeats downstream the IS1E. We conducted a conventional α-complementation analysis by using AAT•ATT triplet repeats free pAAT24IS1E-A and a positive

Figure 3. IS1E transposition into AAT-ATT repeats. A) Plasmids of AAT orientation propagated in LB medium were recovered and digested by PstI (lane 1), EcoRI (lane 2); and propagated in LB medium in the presence of IPTG, and were digested by PstI (lane 3), and EcoRI (lane 4); Plasmids of ATT orientation propagated in LB medium with or without IPTG, digested by PstI (lane 5), EcoRI (lane 6); PstI (lane 7 with IPTG induction) and EcoRI (lane 8 with IPTG induction). M1 and M2 are DNA molecular weights; B) Schematic illustration of the directions of IS1E transposition into the AAT-ATT repeats in light of the transcription of either AAT or ATT orientation. doi:10.1371/journal.pone.0014271.g003
control plasmid, pUC18 by transforming these plasmids into JM83 cells. The transformants were grown on LB plates containing ampicillin, IPTG, and X-gal. JM83 cells carrying the repeats free pAAT24IS1E-A plasmid formed white colonies while the pUC18 plasmid formed blue colonies (Figure 4). Therefore, transposition of IS1E into the upstream region of the AAT•ATT triplet repeats in the lacZ promoter repressed transcription of the repeats. IS1E elements have two partially overlapping open reading frames, InsA and InsB′, which have a relative reading frame of 0 and -1, respectively [29,30,31,32]. InsA protein regulates transcription of the gene producing the IS1E transposase, while InsAB′ binding to both the left and right IS1E terminals represses IS1E transposase transcription from a promoter found partly in IRL and simultaneously inhibits transposition [29,30,31,32]. As shown here, such a regulatory mechanism may contribute to the repression of AAT•ATT triplet repeats transcription.

This has been further confirmed that repression of AAT•ATT triplet repeats transcription by IS1E insertions helped plasmid stability; we co-propagated a LacI2′ producing plasmid with either the pAAT24 or the pTTA24 plasmids. Co-propagation of the LacI2′ plasmid and pAAT24 or pTTA24 improved plasmid recovery, which was also consistent with the result of adding 5% glucose to the medium to repress lacZ gene transcription (data not shown). These results further indicated that transcription of AAT•ATT triplet repeats was responsible for the morphological changes of the E.coli cells, and that inactivation of transcription by IS1E insertions into the lacZ promoter or the AAT•ATT triplet repeats proximal to the promoter helped cells recover from the transcriptional stress.

**H-NS silences transcription by binding to AAT•ATT triplet repeats and promotes IS1E transposition**

H-NS selectively silences bacterial genes associated with pathogenicity, and the gene’s responses to environmental insults [33,34,35,36]. H-NS binds preferentially to the AT-rich motifs displaying planar curvature, which is likely to appear in bacterial promoters [33,36,37]. In addition, H-NS also influences transposition and recombination [38,39,40]. In some transposition processes, strains with hns mutations show a low-level of transposase production. In IS1 transposition, reduction of InsAB′ production to barely detectable levels has been reported [38,39,40]. Because we observed here that H-NS was required for IS1E transposition (Figure 1), while it was also implicated in the colony morphological changes in conjunction with the transposition of the 24 AAT•ATT triplet repeats, we determined whether the hns gene product itself may have a role in triplet repeats transcription by measuring the effects of H-NS on plasmid copy number variations in the wild type and the Δhns mutant of JM83.

The results indicated that transcription of the two orientations of AAT•ATT triplet repeats was differently affected by H-NS. However, the plasmid copy number of pAAT24 was not significantly affected by the presence or absence of H-NS, or by the presence or absence of IPTG induction (Table 1). Only 86% of the plasmid copy number of JM83 was detected in the JM83Δhns mutant when induced by IPTG; 69% of the plasmid copy number of JM83 was detected in JM83Δhns with IPTG induction (Table 1). To rule out that multiple copies of pAAT24 and pTTA24 mitigated H-NS effects in the JM83 wild type, we integrated a single copy of a 24-triplet AAT•ATT repeat in two orientations into the attB site located on the E.coli JM83 chromosome (see Materials and Methods and Figure 1). Propagations of the repeats carrying strains were monitored by analyzing changes of bacterial cell morphology using confocal microscopy. Filamentous growth of the cells was noted when propagating the AAT•ATT repeats in the chromosome of the JM83 Δhns mutant, but not in the JM83 wildtype (Figure 5), nor in JM83Δhns and JM83 strains carrying either the pAAT24 or pTTA24 plasmid (data not shown). These suggested that H-NS affected the AAT•ATT triplet repeats on the chromosome in a dosage dependent manner.

**Discussion**

In human beings, the expansion and contraction of trinucleotide repeats CAG•CTG, CGG•CCG and GAA•TTC are associated with more than 40 human genetic diseases and cancers, including Fragile X, Huntington’s disease, SCA1-12, MD, and Fredericha’s...
ataxia [8,13]. Similarly, as one of the most abundant and most polymorphic trinucleotide repeats in the human genome, the AAT\ATT triplet repeats also show instabilities, which have recently implicated in some human health symptoms, for example, high IgE blood symptom in Chinese children [19], schizophrenia [20,21], cocaine addiction [22], and prevalence of depression in Parkinson’s disease patients [23]. However, understanding the transactions of AAT\ATT triplet repeats in vivo has so far been largely elusive. In this work, we found that a 24-triplet AAT\TTA repeat can be propagated stably in \textit{E.coli} when it was not transcribed. Transcription of the triplet repeats invoked cell toxicity, and therefore had to be silenced by using H-NS or by H-NS facilitated IS1E transposition. In contrast, propagation and transcription of a short AAT\TTA repeat of 8 triplets under otherwise similar conditions did not show cell toxicity, nor H-NS and H-NS facilitated IS1E transposition, suggesting that transcription mediated cell toxicity is dependent on the length of the AAT\TTA triplet repeats, and that similar sized AAT\TTA repeat may also possibly be intrinsically vulnerable for transcription in all eukaryotes and prokaryotes. Consisting with this, we found that AAT\TTA repeats longer than 16 triplets do not exist in the databases of human genomic plus transcript, the mouse genomic plus transcript and the others as set by NCBI. Although we realized that very long AAT\TTA repeats have been found in many species, including human being and Drosophila etc., this may further implicate that transcription acted as selective pressure against long AAT\TTA repeats (>17 triplets) to be distributed in the coding region of genes.

The mechanism underlying the induction of cell toxicity and the cellular morphological changes by transcription of AAT\ATT triplet repeats is complex. In this work, we have ruled out the effects of generation of DNA double strand breaks and the cryptic phage induced cell lysis on colony morphological alterations. While our work directly implicated an effect of transcription of longer AAT\ATT triplet repeats on cellular morphological changes, we reasoned that transcription of longer AAT\ATT triplet repeats may facilitate the repeats to form certain types of non-B DNA secondary structures, such as non-H structure, which recruits binding of histone-like protein such as H-NS etc (Figure 6) [9], causing depletion of nuclear structure associated proteins in the chromosome in cells. The cellular morphological changes could be induced due to the depletion of the histone-like proteins, which may also includes HU and IHF etc in some situations [35,40,41,42]. In support of this idea, it was found that propagation of AAT\TTA triplet repeats containing different triplets formed non-H structure in vivo [9], and also simultaneous depletions of H-NS, HU and IHF in \textit{Escherichia coli} K-12 are lethal [41,42]. Cells under the situation of depletion of histone-like

| Table 1. Effects of H-NS on the plasmid copy number. |
|-----------------------------------------------|
| JM83(AAT) | JM83(ATT) | .hns(AAT) | .hns(ATT) |
| No IPTG | With IPTG | No IPTG | With IPTG | No IPTG | With IPTG | No IPTG | With IPTG |
| Plasmid/Chromo | 0.592 | 0.586 | 0.620 | 0.553 | 0.568 | 0.536 | 0.396 | 0.316 |
| | 0.316 | 0.392 | 0.300 | 0.329 | 0.433 | 0.465 | 0.297 | 0.330 |
| | 0.204 | 0.287 | 0.304 | 0.206 | 0.408 | 0.486 | 0.350 | 0.273 |
| Averaged Ratio | 0.371 | 0.422 | 0.408 | 0.363 | 0.500 | 0.496 | 0.348 | 0.306 |

doi:10.1371/journal.pone.0014271.t001

Figure 5. Effects of \textit{hns} gene on the growth of AAT-ATT repeats-carrying strains in chromosome. Significant filamentous cells were observed in JM83\textit{hns} -AAT and JM83\textit{hns} -ATT, but cannot be seen with JM83 -AAT and JM83 -ATT, nor be seen with JM83(pAAT24), JM83(pATT24), JM83\textit{hns}(pAAT24), JM83\textit{hns}(pATT24) and JM83(pUC18), JM83\textit{hns} (pUC18) (data not shown).
doi:10.1371/journal.pone.0014271.g005
AAT into mixed repeats by IS1E and eventually form “Junk DNA”. The transposition, over time, the repeats might be able to be converted from transcription by using H-NS silenced and H-NS facilitated transposition inactivation, which somehow stabilize the repeats from causing cell morphological alteration and probably also repeats expansion and contraction. While hearing the findings of IS1E transposition into both the promoter region and the repeats array, we are unable to distinguish if the AAT•ATT triplet repeats serve as usual hotspots of transpositions, or does it do so only in the repeats transcription? Our findings of the IS1E transposition depended on the repeats orientation in transcription could suggest that the transposition is linked to the transcription of AAT•ATT triplet repeats. Interestingly, a similar transposition of Tn5 into GAA•ATT repeats during the repeats forming triplex has also recently been demonstrated in vitro [43].

Conclusions

We found that transcriptions of longer AAT•ATT triplet repeats located on plasmid and in the E.coli genome were responsible for the formation of abnormal cell colonies. The bacterial colonies were sick in morphology, which was accompanied by significant IS1E transposition and filamentous cell growth. H-NS was found to be a key protein for these processes; presumably inactivating triplet repeats transcription by differently binding to the repeat arrays and then promoting IS1E transposition (Figure 6). These findings suggest that RNA transcription of AAT•ATT triplet repeats may serve as a selective pressure for biased distribution of AAT•ATT repeats in different chromosomal regions.

Materials and Methods

Bacterial strains

Bacterial strains used in this work were AB1157 recF:: Tn10KanR, W3110△tadD::cat, JM83 wildtype [44], and JM83 derivatives of ΔrevA, recB, Δhns, and ΔhnsrecB, ΔtadD::cat, respectively. Strains JM83-AAT24 and JM83-ATT24 were constructed by integrating the (AAT•ATT)24 repeats at the attB site of the chromosome with two orientations, of which either AAT or ATT strand of the (AAT•ATT)24 served as the template for transcription. Strains JM83Δhns–AAT24 and JM83Δhns–ATT24 were constructed by P1 transduction of the Δhns gene into the chromosome of JM83-AAT24 and JM83-ATT24, respectively. P1 transduction was performed as described in [44].

Plasmids

The plasmids used in this study were pKOV, pUC18 and its AAT•ATT repeats carrying derivatives, pAAT24 and pTTA24 [45,46]. Plasmid pAAT24 was constructed by cloning the (AAT)24 repeats in the EcoRI site of pUC18 plasmid, which was on the lagging-strand template of the replication fork (a gift from C. Abbott, University of Edinburgh). Inversion of the trinucleotide repeat array of (AAT•ATT)24 to generate plasmid pTTA24 was performed by using EcoRI cleavage, and followed by religation using T4 DNA ligase. The plasmids pAAT24 and pTTA24 were all confirmed by DNA sequencing. Plasmids pAAT24IS1E-A, pAAT24IS1E-B, pAAT24IS1E-C, were plasmids of pAAT24 carrying...
IS1E elements in the promoter region of the lacZ (pAAT24IS1E-A, pAAT24IS1E-B) or in the AAT repeat array (pAAT24IS1E-C), respectively. Plasmids pTTA0 and pTTA0 were the deletion products of plasmid pAAT24 when propagated in a JM83 recA::catR mutant. DNA sequencing primer used in this work is 5'-ATCCACATTG GCCCTCCATC-3', which was synthesized by Huada Co, Ltd (Beijing).

Enzymes, antibiotics and biochemicals
Restriction enzymes EcoRI, PstI, NotI used in this work were products of Promega (Beijing); T4 DNA ligase was purchased from New England Biolabs; Ampicillin was from Boehringer Mannheim. Isopropylthio-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) were bought from Sigma Chemical Company, respectively.

Media and bacterial cultivation
Luria-Bertani broth (L-broth) was utilized for the cultivation of bacteria at 30°C and 45°C when constructing the strains of JM83-AAT24, JM83-ATT24 by using pKOV integration [44,47]; while all other types of cultures were carried out at 37°C. Ampicillin was applied by a concentration of 100 μg/ml when it was required. Transformation was performed by using a CaCl2 method [47].

Plasmid DNA isolation and agarose gel electrophoresis
Plasmid DNA was prepared using a kit purchased from Qiagen after the propagations of the plasmid carrying strains for a period of time, normally overnight cultivation was applied. Agarose gel electrophoresis was conducted according the reference [47] on 0.8% gels (Flowgen).

Examination for the repeat tract instability
Plasmids population was examined following the method [28]: Briefly, monomeric plasmid DNA was used to transform E. coli strains of interest and plasmid DNA prepared from a population of transformants. Cells from roughly 4 primary transformants were harvested in 5 ml L-broth and 50 μl of this suspension was diluted into 5 ml L-broth and grown for 24 hours. This corresponds to 30 generations of cell growth. Plasmid DNA was isolated and cleaved with EcoRI. The fragments were end-labelled with S35-dATP using DNA polymerase I Klenow fragment and resolved on 8% native agarose gel. Plasmid DNA was prepared using a kit purchased from Qiagen restriction site (underlined sequence), respectively, and also the JM83-AAT24, JM83-ATT24.

DNA sequencing and determination of the transposon
DNA cycle-sequencing was performed using a kit purchased from PE Applied Biosystems. And DNA sequence was extracted from DNA sequencing primer used in this work were products of Promega (Beijing); T4 DNA ligase was purchased from New England Biolabs; Ampicillin was from Boehringer Mannheim. Isopropylthio-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) were bought from Sigma Chemical Company, respectively.

DNA sequencing and determination of the transposon
DNA cycle-sequencing was performed using a kit purchased from PE Applied Biosystems. And DNA sequence was extracted from DNA sequencing primer used in this work were products of Promega (Beijing); T4 DNA ligase was purchased from New England Biolabs; Ampicillin was from Boehringer Mannheim. Isopropylthio-D-galactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) were bought from Sigma Chemical Company, respectively.

Integration of the AAT-TTA repeats into chromosome attB site
Plasmid pKOV was used as an integrative tool for the construction of JM83-AAT24 and JM83-ATT24 [45]. Two DNA primers were designed and utilized as follows: upper strand primer: 5’ GTGTTTCAGCAGCGCCGTCGGGCGATATGAAA-TAGAAAAATGAAATCCTGTCGCTGGTATACCTAGC, and lower strand primer: 5’ CAGATTGGCGCCCGCACCTGTTGAT-TCATTTAAGGTATTAAACCGCCAGATGCGG3’, which were synthesized by Shanghai Sangon Co. Ltd. They bear a NotI restriction site (underlined sequence), respectively, and also contain part of attB sequence, and the flanking sequence of lacZ open reading frame of plasmids pAAT24 and pTTA24. Integrative plasmids pKOV-AAT and pKOV-ATT were constructed by inserting the PCR products at the NotI sites.

Strains of JM83-AAT24 and JM83-ATT24 were selected against the following criteria [45]. For using antibiotic, 20 mg/ml of chloramphenicol was utilized; for selection against sucB, sucrose was added into the LB medium to a final concentration of 5% (v/v). PCR amplification using the same DNA primers as abovementioned was further performed for the confirmation of the JM83-AAT24, JM83-ATT24.

Observation of the cell morphological changes during cultivation
Morphological alterations of the E. coli vegetative cells grown in LB broth were monitored by using a light microscope with oil immersion objective. Cells were stained by safranin before observation [49].

Acknowledgments
We are grateful to Drs D Leach and C. Abbott at the University of Edinburgh for the AAT·ATT repeats and some strains; Prof. B. Michel in CNRS, France, for the AB1157 recF strain; Dr. Yasuhiko Sekine at Rikkyo University, Japan, for the Abn mutant; Prof. Naotake Ogasawara at Nara Institute of Science and Technology, Japan, for the W3110 attB::cat mutant; Prof. D Lilly at Dundee University for the lacI Q expression plasmid; Prof. George M. Church at Harvard Medical School, USA, for the pKOV plasmid; Prof. M Schlappi at Marquette University for manuscript editing and his constructive comments on the work, And all members of the Pan group for their help and discussion. YHL, YML and PC are research students in the laboratory and contributed equally to this work.

Author Contributions
Conceived and designed the experiments: XP. Performed the experiments: XP YL PC YL LL LY. Analyzed the data: XP HL. Contributed reagents/materials/analysis tools: XP HL. Wrote the paper: XP.
References

1. Astolfi P, Bellizzi D, Sgaramella V (2003) Frequency and coverage of trinucleotide repeats in eukaryotes. Gene 317: 117–125.
2. Subramanian S, Mishra RK, Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Bio. 8: R13.
3. Rocha EFC, Matie I, Taidelli F (2002) Over-representation of repeats in stress response genes: a strategy to increase versatility under stressful conditions? Nucleic Acid Res 30: 1886–1894.
4. Gur-Arie R, Cohen CJ, Eitan Y, Shelef E, Halleman EM, et al. (2000) Simple sequence repeats in Escherichia coli: abundance, distribution, composition, and polymorphism. Genome Res 10: 62–71.
5. Toth G, Gaspari Z, Jurka J (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10: 967–981.
6. Cox R, Mirkin SM (1997) Characteristic enrichment of DNA repeats in different genomes. Proc Natl Acad Sci USA 94: 5237–5242.
7. Todd J, Treangen TJ, Abraham A-L, Touchon M, Rocha EPC (2009) Genesis, effects and fates of repeats in prokaryotic genomes. FUNEM Microbiol Rev 1–33.
8. Everett CM, Wood NW (2004) Trinucleotide repeats and neurodegenerative disease. Brain 127: 2385–2405.
9. Ohshima K, Kang S, Larson JE, Wells RD (1996) TTA repeat non-B secondary structure formation and the roles of cellular trans-acting factors. J Genet Genomics 33: 1–11.
10. Ona E, Del Grosso N, Erba M, Pari M, et al. (2006) Interaction of DAPI with individual strands of trinucleotide repeats adopts stable hairpin structures induced by minor groove binding ligands. Biochemistry 39: 6799–6808.
11. Dayn A, Samadashwily GM, Mirkin SM (1992) Intramolecular DNA triplexes: Unusual sequence requirements and influence on DNA polymerization. Proc Natl Acad Sci USA 89: 11406–11410.
12. Oshima T, Ishikawa K, Kurakawa K, Alba H, Ogawa N (2006) Escherichia coli histone-like protein H-NF preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res 13: 141–153.
13. Pan X (2006) Mechanism of trinucleotide repeats instabilities: the necessities of repeat non-B secondary structure formation and the roles of cellular trans-acting factors. J Genet Genomics 33: 1–11.
14. Otta E, Del Grosso N, Erba M, Melino S, Cicero D, et al. (2003) Interaction of DAPI with individual strands of trinucleotide repeats: effects of replication in vitro of the AAT · ATT triplet. Eur J Biochem 270: 4755–4761.
15. Hubacek JA, Pistulkova H, Valenta Z, Polelde R (1999) TTA repeat polymorphism in the HMG-CoA reductase gene and cholesterolemia. Vasa 28: 169–171.
16. Lysons-Darden T, Topal MD (1999) Effects of temperature, Mg2+ concentration and mismatches on triplet-repeat expansion during DNA replication in vitro. Nucleic Acids Res 27: 2235–2240.
17. Kappen LS, Xi Z, Jones GB, Goldberg IH (2003) Stimulation of DNA strand slippage synthesis by a bulge binding synthetic agent. Biochemistry 42: 7975–7980.
18. Leung TF, Liu EK, Tang NL, Ko FW, Li CY, et al. (2005) Nitric oxide synthase CNR1: association with schizophrenia in a Spanish population. Eur Arch Psychiatry 256: 1053–1057.
19. Ujike H, Takaki M, Nakata K, Alba H, Ogawa N (2006) Escherichia coli histone-like protein H-NF preferentially binds to horizontally acquired DNA in association with RNA polymerase. DNA Res 13: 141–153.
20. Rouquette C, Serre M-C, Lane D (2004) Protective Role for H-NS protein in vivo of the AAT · ATT triplet. J Biol Chem 270: 4755–4761.
21. Amici R, Oppenheim AB, Stavans J (2003) Increased Bending Rigidity of Single DNA Molecules by H-NS, a Temperature and Osomolarity Sensor. Bioch J 384: 231–234.
22. Briol SW, Berg DE (1984) mechanism of IS1 transposition in E.coli TTA triplet repeats in plasmids form a non-B structure. J Biol Chem 217: 1679–1679.
23. Trotta E, Del Grosso N, Ebra M, Pari M (2000) The ATT strand of AAT/ATT trinucleotide repeats adopts hairpin structures induced by minor groove binding ligands. Biochemistry 39: 6799–6808.
24. Sawada T, Ueda H, Nishio K, et al. (2002) Is the IS1 transposase, InsAB', the only IS1-encoded protein required for efficient tranposition? J Bacteriol 176: 5964–5967.
25. Schwartz M, Zlotorynski E, Kerem B (2006) The molecular basis of common and rare fragile sites. Cancer Lett 232: 15–26.
26. Pan X, Ding YF, Shi LF (2009) The roles of ShcCD and RbXase in the transcription of GAA/TTG repeats in Escherichia coli. DNA Repair (Amst) 8: 1321–1327.
27. Komiarda B, Wells RD (2006) Double-strand breaks in the myostatin dystrophy type 1 and the fragile X syndrome triplet repeat sequences induce different types of mutation in DNA flanking sequences in Escherichia coli. Nucleic Acids Res 34: 5309–5312.
28. Pan X, Leach D (2000) The roles of mutS, bcd/pcb and reh in the propagation of TGG repeat in Escherichia coli. Nucleic Acids Res 28: 5378–5384.
29. Escobar JM, Lane D, Chaitelot M (1994) Is the IS1 transposable, InsA?, the only IS1-encoded protein required for efficient transposition? J Bacteriol 176: 5964–5967.
30. Lane D, Cavaille-J, Chandler M (1994) Induction of the SOS response by IS1 transposase. J Mol Biol 241: 357–372.
31. Matsutani S (1997) Genetic Analyses of the Interactions of the IS1-encoded proteins with the left end of IS1 and its insertion hotspot. J Mol Biol 267: 540–560.
32. Biel SW, Berg DE (1984) mechanism of IS1 transposition in E. coli: Choice between simple insertion and cointegration. Genetics 108: 319–330.
33. Fang FC, Rimsby S (2008) New insights into transcriptional regulation by H-NS. Curr Opin Microbiol 11: 115–120.
34. Durna RT, Wymann C, Grosse M, N (2003) Structural basis for preferential binding of H-NS to curved DNA. Biochimie 83: 231–234.
35. Amit R, Oppenheim AB, Stavans J (2003) Increased Bending Rigidity of Single DNA Molecules by H-NS, a Temperature and Osomolarity Sensor. Bioch J 384: 231–234.
36. Shiga YY, Seikine YK, Ohtsubo E (2001) Involvement of H-NS in transposition regulation mediated by IS1. J Bacteriol 183: 2476–2484.
37. Bouffartigues E, Buckle M, Geertz M, et al. (2007) High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within protobacterial genomes. Nucleic Acids Res 35: 6330–6337.
38. Pan X, Ding YF, Shi LF (2009) The roles of ShcCD and RbXase in the transcription of GAA/TTG repeats in Escherichia coli. DNA Repair (Amst) 8: 1321–1327.
39. Escobar JM, Lane D, Chaitelot M (1994) Is the IS1 transposable, InsA?, the only IS1-encoded protein required for efficient transposition? J Bacteriol 176: 5964–5967.
40. Rouquette C, Serre M-C, Lane D (2004) Protective Role for H-NS protein in vivo of the AAT · ATT triplet. J Biol Chem 270: 4755–4761.
41. Amit R, Oppenheim AB, Stavans J (2003) Increased Bending Rigidity of Single DNA Molecules by H-NS, a Temperature and Osomolarity Sensor. Bioch J 384: 231–234.
42. Shiga YY, Seikine YK, Ohtsubo E (2001) Involvement of H-NS in transpositional regulation mediated by IS1. J Bacteriol 183: 2476–2484.
43. Rouquette C, Serre M-C, Lane D (2004) Protective Role for H-NS protein in vivo of the AAT · ATT triplet. J Bacteriol 176: 2091–2098.
44. Miller JH (1992) A short course in bacterial genetics. N.Y: Cold Spring Harbor Laboratory, Cold Spring Harbor.
45. Phillips AJ, Church GM (1997) Methods for generating precise deletions and insertions in the genome of wild-type Escherichia coli. Application to open reading frame characterization. J Bacteriol 179: 6228–6237.
46. Yamisch-Perron C, Vierra J, Mesjing J (1983) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103–119.
47. Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, Cold Spring Harbor Laboratory.
48. Chen J, Zhang GC, Xu Y, Lu T, Wang CF (2001) Stability and copy number in host bacteria of a new plasmid cloning vector. Biotechnology 11: 6–8.
49. Pan X (2006) Effects of degU32(Hy), degQa and degR Pleiotropic Regulatory Genes on the Growth and Protease Formation of Bacillus subtilis Ki-2-132. J Genet Genomics 33: 373–380.