Repetitive sequences in the human genome

This article has been translated from WikiSkripta; ready for the **editor's review**.

The human genome consists of non-repetitive or unique DNA and from repetitive DNA sequences. The repetitive sequences are important both in gene expression and, for example, in determining the architecture of chromosomes or the cell nucleolus.[1]

Distribution

If the repetitive sequences are in a row of consecutive blocks, periodically repeating sequences, then these are **tandem** repeat sequences. In contrast, **dispersed** sequences are randomly scattered throughout the genome.

Tandem repeats

VNTR (Variable number of tandem repeats) – Polymorphism detectable by the Southern method, specific probe, or PCR. Each copy follows one another on a specific location in the genome. They are variable in the number of repeated units on the designated locus between different chromosomes.

- It is divided into subgroups according to the length of the repetitive sequence:
  1. **macrosatellite**
     - especially around centromeres
     - cztogenetically detectable as regions of heterochromatin (especially chromosomes 1, 9,16, and Y)
  2. **minisateliite**
     - repetitive unit of 5–30 pb
     - underlying multiple polymophisms
  3. **microsatellite**
     - underlying multiple DNA polymorphisms as well (used for indirect diagnosis, person identification)
     - regions consisting of only one nucleotide (especially (A)n) exist as well
  4. **telomeric repeas**
     - multiple kb long regions *(TTAGGG)n*
     - shortened during cell division, in some cells (germinal, stem, some tumor cells), it is maintained by telomerase.

Dispersed

Dispersed repeats are most often created by a process of **transposition, the 'jumping' of a DNA sequence to another place in the genome.** Most of those sequences have the ability to move within the DNA. The sequences are then called transposons.

1. **Retrotransposones** - their typical characteristics is the ability to 'copy' themselves. This process is initiated by the transcription of DNA into RNA and the isertion of the segment back into the original strand by reverse transcriptase. The sequence type **LINE** - *long interspersed nuclear elements* (5 – 7 kb) is autonomous. Thus, it encodes proteins necessary for retrotransposition. The second type of sequence is **SINE** - *short interspersed nuclear elements*. Those are much shorter (100-400 bp). They tend to be associated LINE sequences that enable their mobility
2. **DNA transposons** - in this case, the sequence is 'cut out' and moved to another area of the genome.

From the immediate point of view, transposons do not have any important function in the cell - they are called junk DNA; or selfish DNA, as transposons are propagated at the expense of the cell's energy resources. From a broader point of view, retrotransposon mobility may be important for genome plasticity.[2]

Importance

They contibute to **genome plasticity**, which is understood to be an important **evolutionary** benefit. However, they can be a sourse of **non-homologous recombinations** leading to deletions, inversions, etc. At an indvidual's level, they can also be a source of harmful **mutations**. Amplification of an intragenic repetitive sequence (mostly trinucleotide) can result in a group of human hereditary diseases.

Syndromes caused by trinucleotide repeat expansions

```
Iron
```

Syndromes caused by trinucleotide repeat expansions are cause by unstable/dynamic mutations. There are inherited changes in the number of repeats of three nucleotides (*CAG, CTG, CGC, GAA*) inside or outside the respecitve gene. The disease does not manifest itself until a certain number of trinucleotide repetitions, we are

talking about a so-called **premutation.** During the process of meiotic division, the numbers of repeats can increase up to a **full on mutation.** The full mutation is then passed in to an affected offspring in the next generation.

Examples of syndrome syndromů

thumb|For the Fraxile X syndrome, an elongated face, large ears, and prominent chin are tzpical.|200px

- **Huntington's disease**
  - 5´end of a gene – repetition of *CAG*
  - naturally there are 10–34 repetitions
  - the affected have 42–100 copies
- **Myotonic dystrophy**
  - repetition of the *CTG triplet* after the 3´ end of a gene
  - the affected have50 repetitions or more
- **Friedreich's ataxia**
  - *GAA* repetition in the gene intron
- **Fragile X syndrome (fraX)**
  - Amplification of *CCG* triplets in the *FMR1* gene in the *FRAXA* (fragile site) region, which is located on the long arms of the X chromosome.
  - Up to 50 repeats – normal allele
  - 50–200 repetitions – premutation, usually without a phenotypic manifestation
  - 200–230 repeats– full mutation

References

Related articles

- Huntington's disease
- Fragile X syndrome

Used literature

- KOČÁREK, Eduard – PÁNEK, Martin – NOVOTNÁ, Drahuše. *Klinická cytogenetika I : úvod do klinické cytogenetiky, vyšetřovací metody v klinické cytogenetice.* 1. edition. Karolinum, 2006. 120 pp. ISBN 80-246-1069-8.

- GOETZ, Petr. *Vybrané kapitoly z lékařské biologie I.* 1. edition. H & H, 1994. 176 pp. ISBN 80-85787-56-3.

- GOETZ, Petr. *Vybrané kapitoly z lékařské biologie II.* 1. edition. Karolinum, 2002. 139 pp. ISBN 80-246-0320-9.

- MOYZIS, R K – ALBRIGHT, K L – BARTHOLDI, M F. Human chromosome-specific repetitive DNA sequences: novel markers for genetic analysis. *Chromosoma* [online]. 1987, vol. 95, no. 6, p. 375-86, Available from <https://www.ncbi.nlm.nih.gov/pubmed/3677921>. ISSN 0009-5915.

References

1. MOYZIS, R K, K L ALBRIGHT a M F BARTHOLDI, et al. Human chromosome-specific repetitive DNA sequences: novel markers for genetic analysis. Chromosoma [online]. 1987, vol. 95, no. 6, s. 375-86, dostupné také z <https://www.ncbi.nlm.nih.gov/pubmed/3677921>. ISSN 0009-5915.
2.

Category:Genetics Category:Molecular biology

2. https://biol.lf1.cuni.cz/ucebnice/repetitivni_dna.htm