

# Hesaplamaalı Biyoloji ve Biyoinformatiğe Giriş



Teorik: Barış Özudoğru  
Pratik: Işın Altinkaya

# Biyoinformatik nedir?

*Biyoinformatik*, biyolojik araştırmayı hızlandırmak ve geliştirmek için bilgisayar veritabanları ve algoritmalar geliştirme bilimidir.

# Hesaplamaalı Biyoloji Nedir?

Geniř Anlamda, bilgisayar bilimleri, istatistik ve matematięin biyolojik bilimlerdeki her türlü probleme uygulanmasıdır.

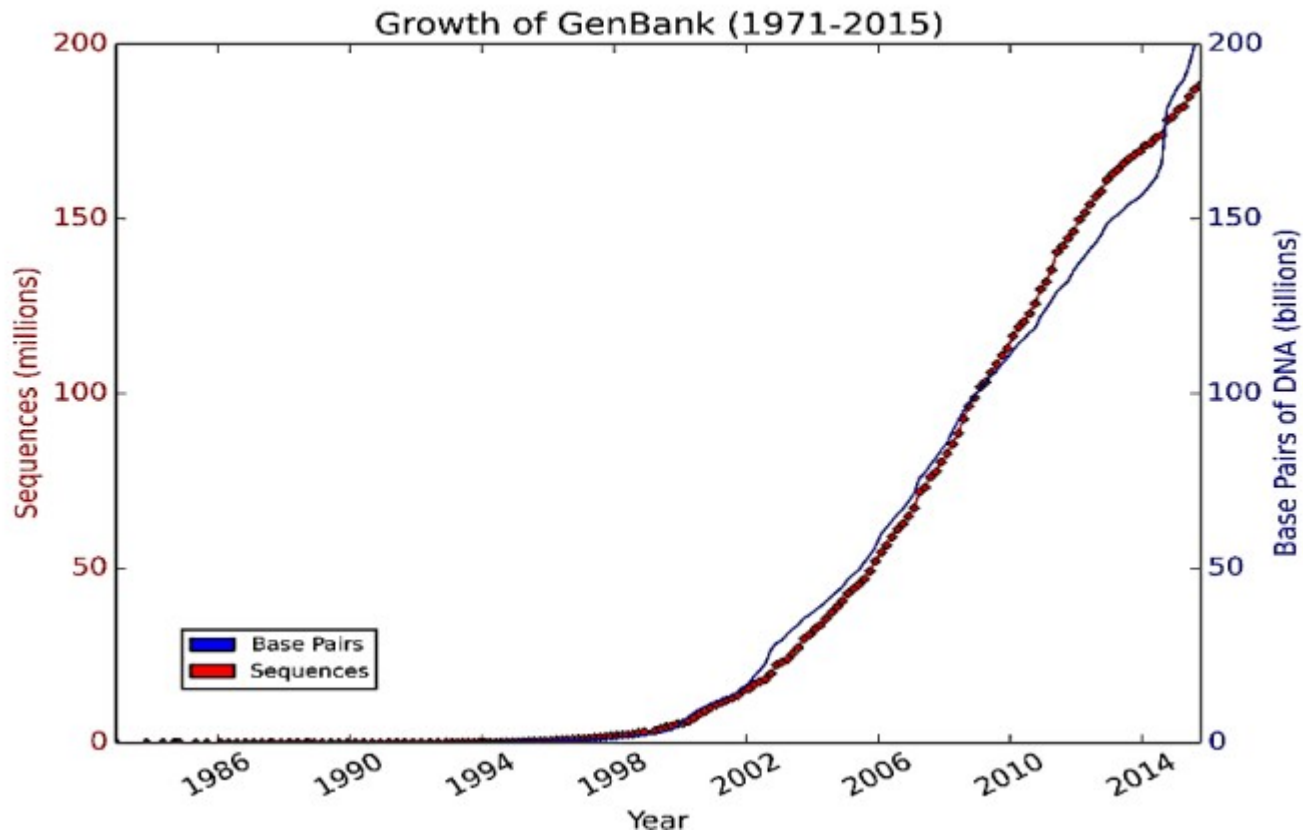
Hesaplamaalı biyoloji, biyoloji içinde genomik / genetik, biyofizik, hücre biyolojisi, biyokimya ve evrim de dahil olmak üzere çok çeřitli alanları kapsar.

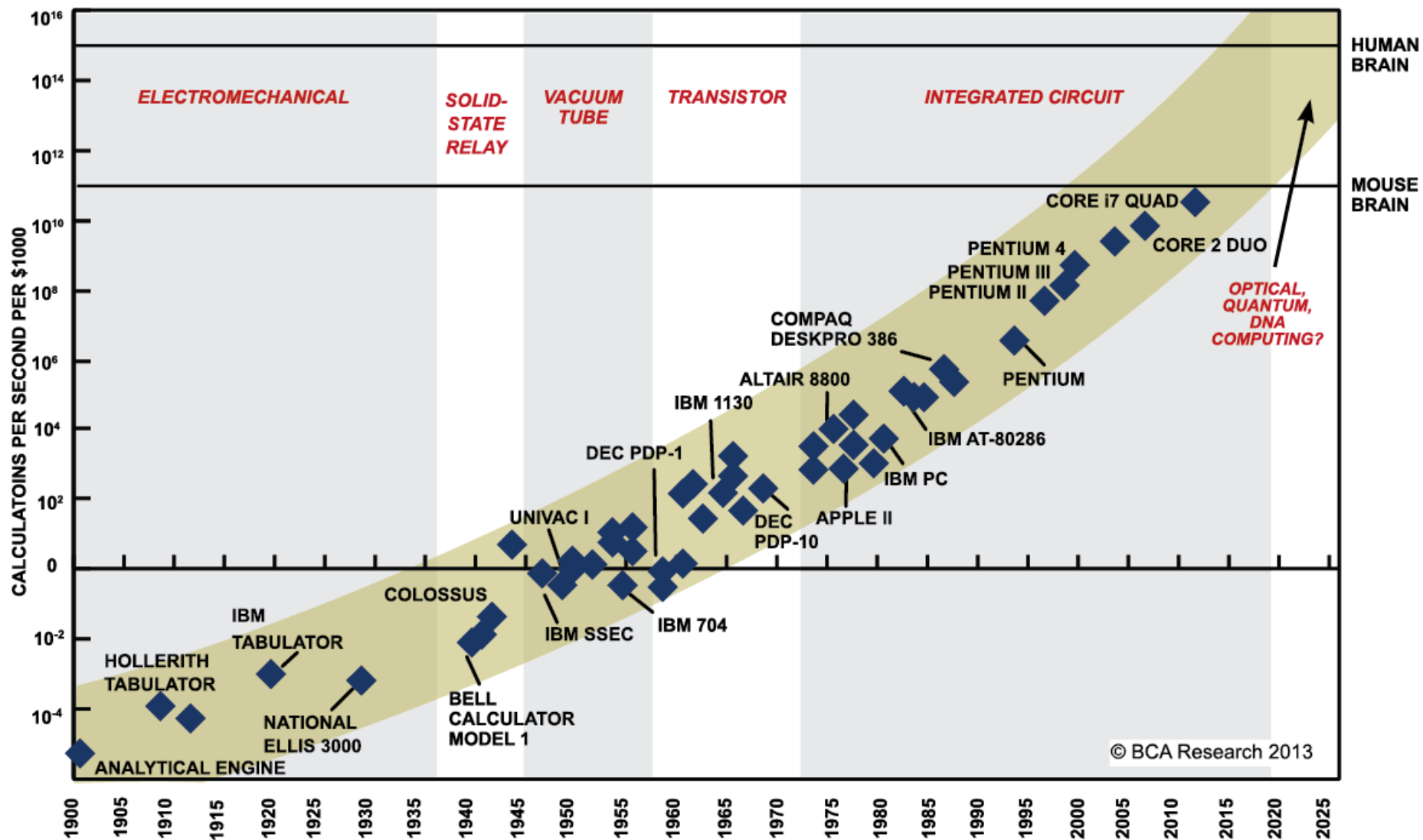
# Hesaplamalı Biyolog ne çeşit problemlerle uğraşır?

- Biyolojik sekansların (DNA, RNA, protein) analizi,
- 3D protein yapıları,
- Moleküler biyolojik networkler (metabolik yollar, protein-protein etkileşimi)
- Klinik ya da ekolojik veriler

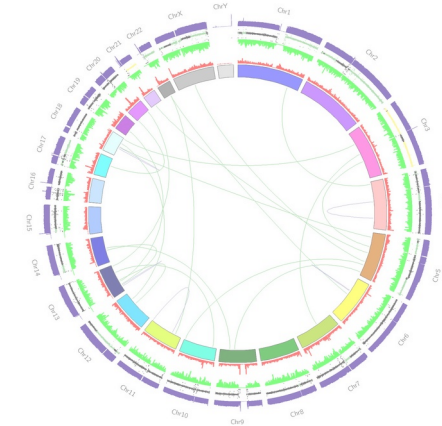
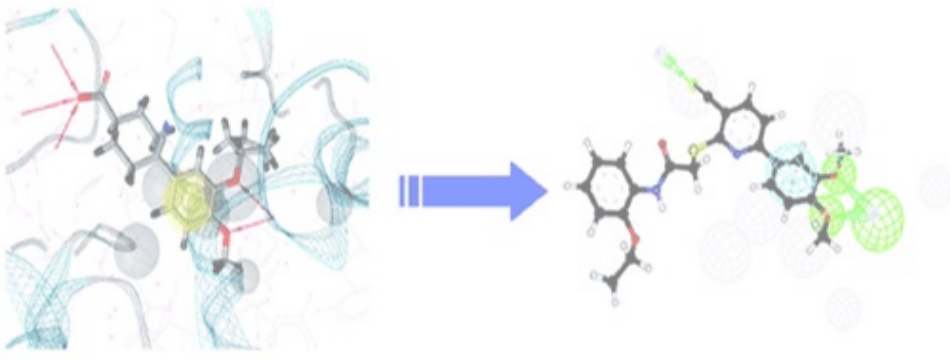
# Hesaplamalı Biyoloji vs. Biyoinformatik

# Hesaplamaalı Biyoloji'ye Giriş





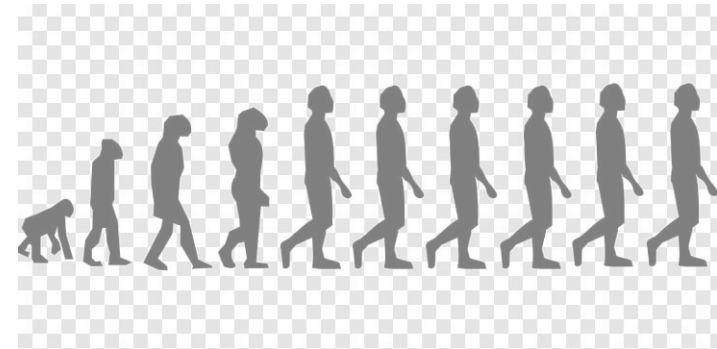
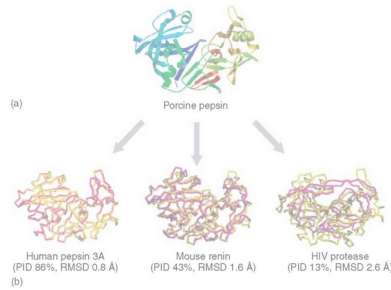
SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, THE VIKING PRESS, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.



İlaç tasarımı

Genom Analizleri

Hesaplama Biyoloji



Veri tabanları

3B modelleme

Evrimsel Antropoloji

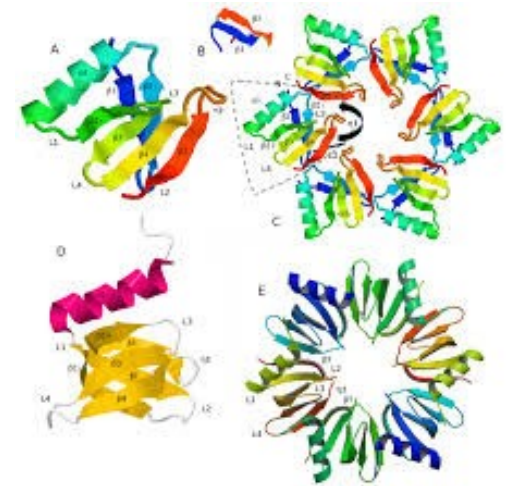


# Biyolojik Veri Tipleri

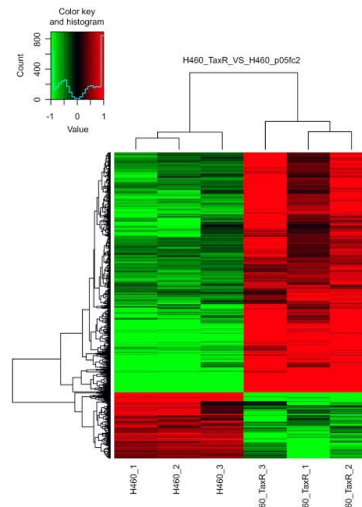
**Diziler:** DNA, RNA, protein

```
170      180      190
ATCTCTTGGCTCCAGCATCGATGAAGAAGCA
TCATTTAGAGGAAGTAAAAGTCGTAACAAGGT
GAACGTGTCAAAACTTTAAACAACGGATCTCTT
TGTTCGTTTCGGCGGGCGCCCGCAAGGGTGCCCG
GGCCTGCCGTGGCAGATCCCCAACGCCGGGCC
TCTCTTGGCTCCAGCATCGATGAAGAAGCAG
CAGCATCGATGAAGAAGCAGCGAAACGCGAT
CGATACCTCTGAGTGTTCTTAGCGAACTGTCA
CGGATCTCTTGGCTCCAGCATCGATGAAGAAG
ACAACGGATCTCTTGGCTCCAGCATCGATGAA
CGGATCTCTTGGCTCCAGCATCGATGAAGAAG
GATGAAGAAGCAGCGAAACGCGATATGTAAT
```

**Biyolojik moleküllerin sekonder ya da tersiyer yapıları**

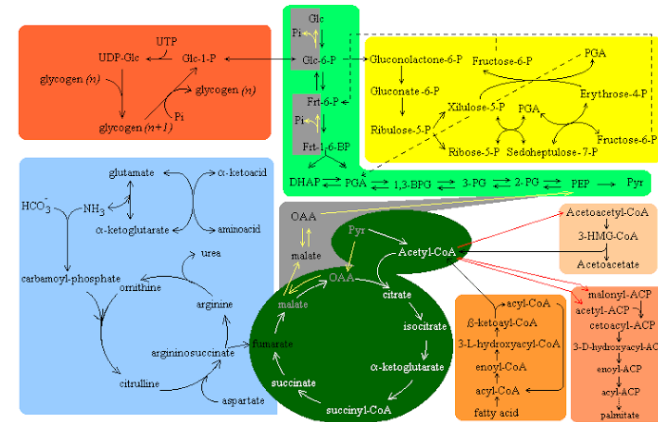


**Gen ifadesi profilleri**

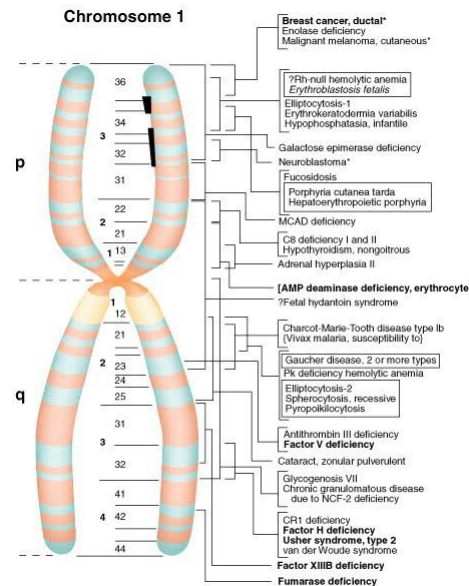


# Biyolojik Veri Tipleri

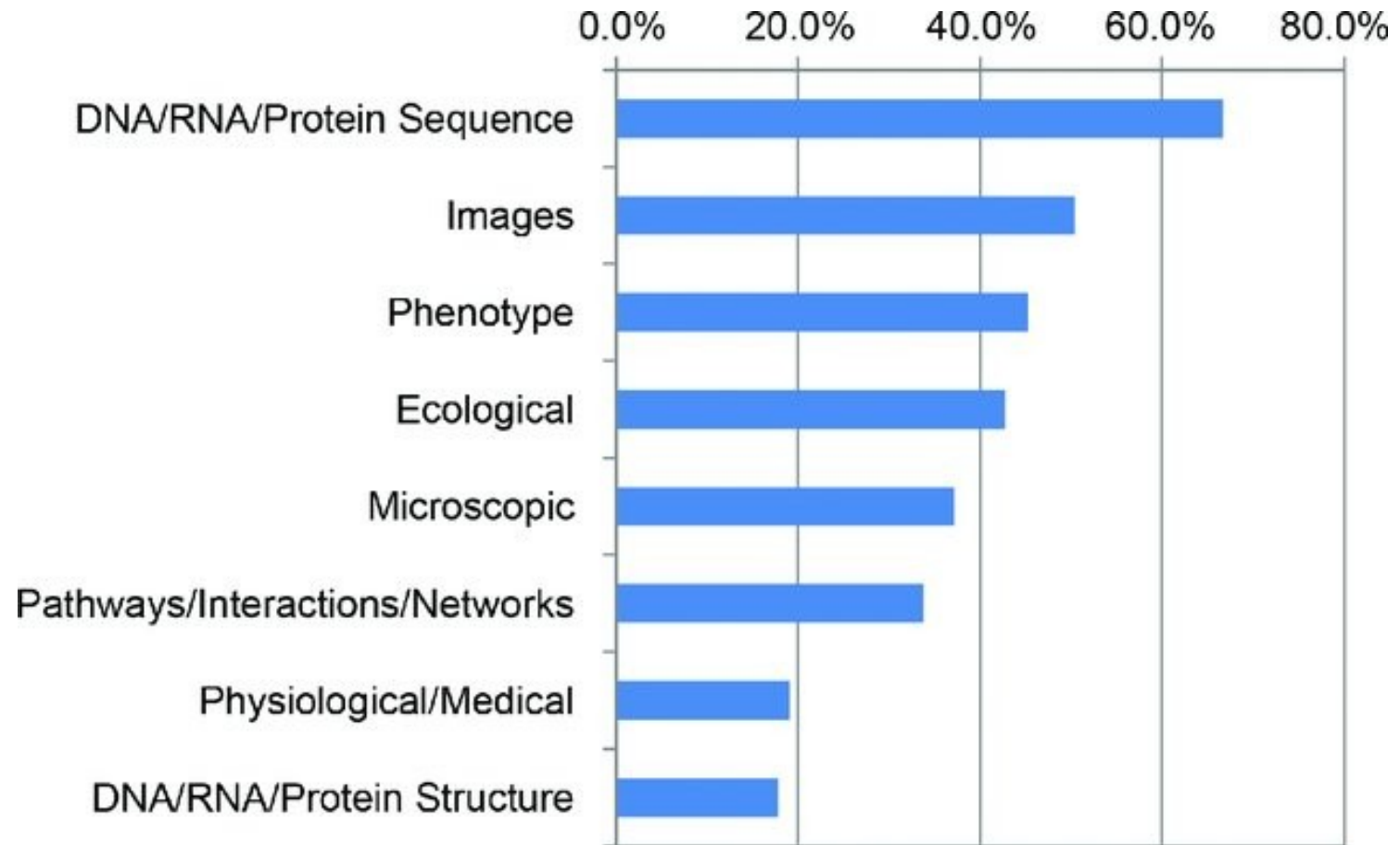
## Biyokimyasal yollar



## Kromozom haritaları



# Biyolojik Veri Tipleri

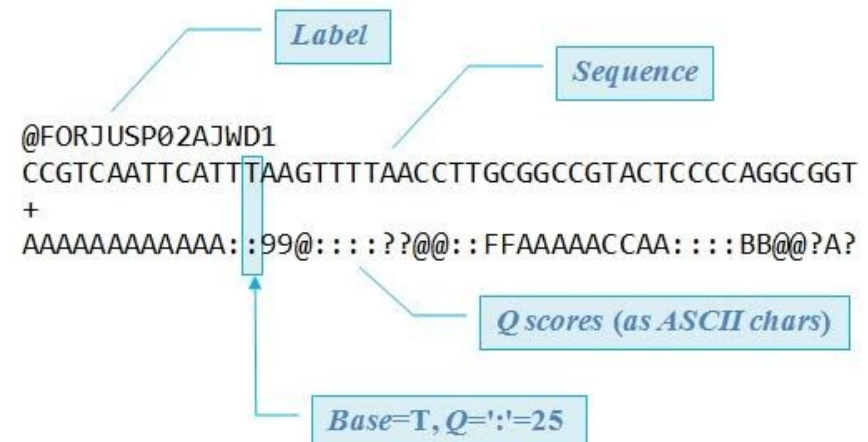


Major data types used by National Science Foundation (NSF) Biological Sciences Directorate (BIO) principal investigators.

# Dosya Formatları

>XP\_210035 loc=GXL\_175098|sym=FAM149A|taxid=9606|spec=Homo sapiens|chr=4|ctg=NC\_000004|str=(+)|start=187065495|end=187066181|len=687|comm=Promoter Region

GGACGGGCGTGGGAAGGGTCCACGCTCTTTAGTATGCATGCTTAGATCTAGCGTTCTCTGTTGATGGAGTAATGGTTCTCGCA  
TTGACCAGATCCGGGGCTTCATTTTTTAAACCTCATTGCTCCACTCCCCACCCAGCCTGGTGTGCGCACCCCTTTGATGG  
GGCGGGGATAGGCGAGATGGTCTGTGTTCTCTGCCTTCTTCTGGTGAATTTAAATCCGATTTGGAAGAGAGAAGGGCA  
GCCAGCACCAAGTATGCACAGCCCCCGCCCCAGAGACCCGGGAAGGAGTAGGGAGGCCGGGCCGTGCGCGGAGGAGTGGC  
CGCTGGGTTGGAAACCCGGCCCGGCAGGGAGCGGGGAAGGCGCGCTTTCCTGGAGGTTCGGCGCGGGGCCGGGGCCGGGGC  
CGGGGCCCGGAGCGGGGATGGGCGGGCGCAGCCGGGATTAGCTGGCGGGCGAGGGCGCAGCGCAGGGAGGAGAGGGGAG  
GCGGCGCCGGCGCGGGCGGGGCGGAGGATTGAGAGAGGAAGGGGCGTGCAGACCCCGCGGACCCGGGGCGCGCCCGGGC  
CGCTTGAGCTGGGCGCAGTCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGGCGGGTGGGGAGCCCCAGCCCC  
GGGGCGCGGGGGCGCGTACCGCGCTGTCTGCGTGGGGGCGCGCGC

[illegible]

# Referans Genom (FASTA)

```
>CP002684.1 Arabidopsis thaliana chromosome 1 sequence
CCCTAAACCCCTAAACCCCTAAACCCCTAAACCTCTGAATCCTTAATCCCTAAATCCCTAAATCCTTTAAATCC
TACATCCATGAATCCCTAAATACCTAATTCCCTAAACCCGAAACCGGTTTCTCTGGTTGAAAATCATTGT
GTATATAATGATAATTTTATCGTTTTTATGTAATTGCTTATTGTTGTGTGTAGATTTTTTAAAAATATCA
TTTGAGGTCAATACAAATCCTATTTCTTGTGGTTTTCTTTCTTCACTTAGCTATGGATGGTTTTATCTTC
ATTTGTTATATTGGATACAAGCTTTGCTACGATCTACATTTGGGAATGTGAGTCTCTTATTGTAACCTTA
GGGTTGGTTTATCTCAAGAATCTTATTAATTGTTTGGACTGTTTATGTTTGGACATTTATTGTCATTCTT
ACTCCTTTGTGGAAATGTTTGTCTATCAATTTATCTTTTGTGGGAAAATTATTTAGTTGTAGGGATGAA
GTCTTTCTTCGTTGTTGTTACGCTTGTCTCATCTCTCAATGATATGGGATGGTCCTTTAGCATTAT
```

## Okumalar (reads) (FASTQ)

```
@HWUSI-EAS100R:6:73:941:1973#0/1
##
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAA
+
!"*(((***+))%%%%%%%%)(%%%%%%%%).1***-+*)"**55CCF>>>>>CCCCCCC65
```

BAM ve BAI

## SAM

```
readID43GYAX15:7:1:1202:19894/1 256
contig43 613960 1 65M * 0 0
CCAGCGCGAACGAAATCCGCATGCGTCT
GGTCGTTGCACGGAACGGCGGCGGTGTG
ATGCACGGC EDDEEDEE=EE?DE??
DDDBADEBEFFFD BEFFEBBCBC=?
BEEEE@=:?:??:8-6??:@??# AS:i:0 XS:i:0
XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0
MD:Z:65 YT:Z:UU
```

## FASTA

```
>N_ibe_beydag_4504_04
TACCTTAGTATGGCCTTTGATTGTGGCAACGTGC
AATGGCGTGTTCCTTTATTATTCTC
AACACTCAAAACAGCAGGATCAGGTTTCACCAA
CTCAAGAACGATCCCTTCGTTTTGACC
TTTGACAGCCATGTGAAGAGCTGTTTGTCCCTTC
ATATCGGTTCTAAACCCGATGCTTGA
ATCGTTACCAATCAAAGACTTNNNNNNNNNNNN
NNNNNNNNNNNNNGCTGCAGAATGAAG
>N_ibe_eldiva_3941_04
TACCTTAGTACGGCCTTTGATTGTGGCAATGTGC
AATGGCGTGTTCCTTTATTATTCTC
AACACTCAAAACAGCAGGATCAGGTTTCACCAA
CTCAAGAACGATCCCTTCGTTTTGACC
TTTGACAGCCATGTGAAGCGCTGTTTGTCCCTTC
TTGTCGGTTCTAAACCCGATGCTTGC
ATCGTTACCAATCAAAGACTTNNNN
```

# FASTQ

Sequence ID  
=dizi kimliği

Sequence  
=Dizi

@HWUSI-EAS100R:6:73:941:1973#0/1

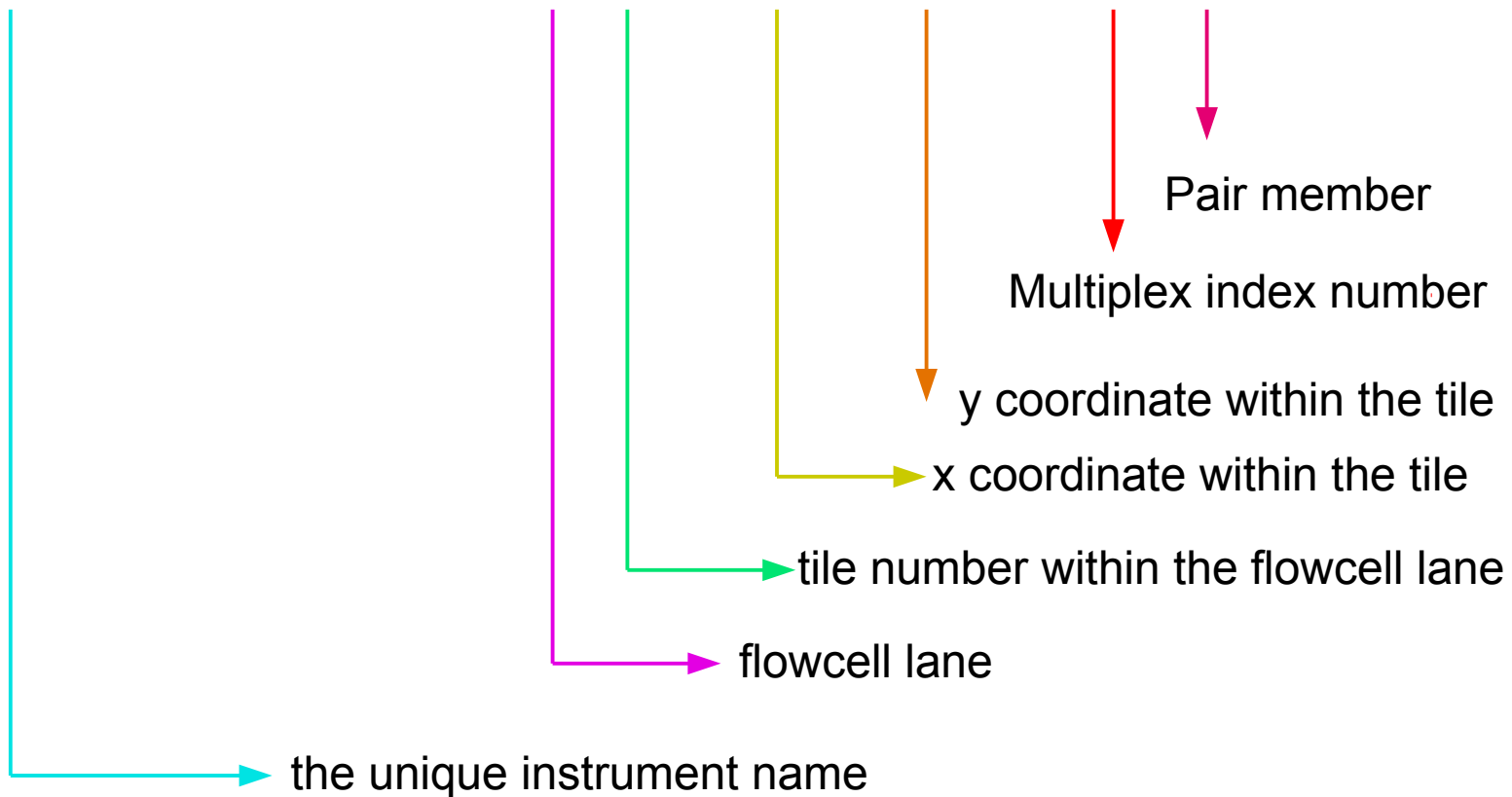
## GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCACAGTTT  
+

!"\*((( (\*\*+))%%%++)(%%%%).1\*\*\*-+\*))\*\*55CCF>>>>>CCCCCCCC65

Kalite skorları

# Sequence ID: Headers

@HWUSI-EAS100R:6:73:941:1973#0/1





# Kalite Skorları



```
S - Sanger          Phred+33, raw reads typically (0, 40)
X - Solexa          Solexa+64, raw reads typically (-5, 40)
I - Illumina 1.3+   Phred+64, raw reads typically (0, 40)
J - Illumina 1.5+   Phred+64, raw reads typically (3, 40)
                    with 0=unused, 1=unused, 2=Read Segment Quality Control Indicator (bold)
                    (Note: See discussion above).
L - Illumina 1.8+   Phred+33, raw reads typically (0, 41)
```



# SAM

QNAME

# CIGAR

## DNA dizisi

## Kalite skoru

[illegible]

# BAM

SAM formatının ikili (binary) formata dönüştürülerek sıkıştırılmış versiyonudur.

<B>>"H'D'@e"@e@FF">F^@BC'B'e="AU<90><CB>J<C3>@T<86>0<84>"Sj">7">@<8D>sM2Y<A5>~W<D0>J<8C><B5>~U7.BLS^Rh:6<B5><BE><86><E0>[K]00^cQ^X<86><EF><FB><CF><C0><F9>gn<D4>^f^Z<80>s>B<B7><9E>Mt<86>fwv<A8>T60VA^Qu<9D><D9>S<9A>y<F6>0<CE>^Hr=<9B>aNM<C3>\$-0+0<A4>!-jdq<9C><9B>V^S<A6>I<A1><C2>X<AD><E6>Y^R<D4><CB>^Y<86><A4><BC><EB>L<CE><D0>xd<DF>?^fh<E2><D5>W<D9>^L<EB><A6>N<E8>IF0<93>h<E8>V<BA><97><FF><A4><F4>7<ED><EE>^B@<A7>X<D5>0<C1><E1>N<F6>^g^T<D6><CF><C7>^M0<F8><FA><8E>^Z<E0><F0><F9>Fd<80><83><F7>7m<AF><84>VeX<DD><ED><C0>^@<83>44<90><F4><D8>U<D8><FB><DB><E7>[]<95>,^ka^yH<A9><A4>J5<FB>-<E0>+n<AC>4w<FC>X<B3>a^Gx1<C0>66^d^R^BNXW<9C><A4>^RVB<85>^E<E1><90>T^Qq<94>ua^Mu^UB<E7><DC>i<F5><EA><96>nw<CF>h^XA3<9D>q<BA>G<AD><EE><D9><F9><E9><DC><F3><BA><E7><9E><EB><DE>2<90><E3>7<8C><C5><E6>q<BA><F4>^<9F><A6><CF>X<BB>^M<B6><9E>~^~^q<B6><FE>[]<D8>x<8B><EA><9E><C5><E2>c<F0>q<E1>]<8B><85>jHj<9D><C8>bBH<9E><D7><F0><A5><B7><97>^B<((<B6><CE>N<E8><F8><DD><D6><F0>qM<D0>5]<A1>^W<E1><C9>[]N<9E>k<E4>^U[k]^I<B0><E1>n<8D>K<D5><ED><DB>^B<BC>I<EB>6^\\<BA><AD><91>6.;:<8E>/)^<E6>3<9F>9<91>6<E0><A3>}<FC><E0>b<E1>t<87>!Ke!fe<AD>:emZ9<95><E5>T<8E>%^W<96><EA>^X@<B6><C8>-<D1>q<8A>:SE<B1><94>d<B1>^E<FC><D3><E1>]

J<83>>7<9F>^<82><8C><97><BA>^"@^\_rh<C0><AB>h<DO>[]^Er^Y]^<CB>^]Y^vt<9D><BC>[]<D7>ESC<8A>k<A4><AB>^AE<C7>(<FE>^W<E6>^P<C9><EB>(<E2<F8>@<F8>dI4<AB><AA><86><AF><A2><AE><AF>@<D8><EE>!<82><EF><DA><C8>[]<A8>^J<9C><BA><B9>C^\_<A2>^!<CA>W^<CC>Wt^g<((<BB>A<8C><E7>^==<8E><86><E6><F9><C1>V<B3>[]<FD><AD>^W<AD><83>^<F1>^R<B0><D2><EF>J,<9E><83><B7><D6>Ta^<9A>\$<B9>(gyU^00<E1>3<96><93>[]a<AA><A5><94><AB><85>\*<AB><95><9A>Iu<A1>fn-<BA>U^B<DF>-S<AC>r<EB><F9>k1<9A><8F>80i^@<98>U<F1>,<D8><F9>I<B8>sw<EE><A5>G^F<B0>^L^@ESCH<9D><98>^C<A4><C2>^Ah<B5>^A<BA>N-[]^RK^M^')<E5><BA>4<C4>]^q<93>D.<D3>Tjr+<CF>J<C3>H<D4><C4>H\$Q^Ve<A3><94>^Y<93>]<A9>^Y^E\$<E8><h<B5><F5>^Bu^S-<9E><83><F1><EB><C2><E7>]<FC><FE><C5>B<94>A<E2><8C>2<AD>^Lxt<A9>:]<A2>qv<C0>^S<8E><D2><C8><E1>,<)<CF><EA>T^U<8B>^R<C<82>j<8A>^ZZP^JB<AC><84>.<88>q!<E7><D0>n<A0><B9>^Da+<BA>[]<8E><E6>ESC^XceylQ<D2>KQ<8F>K/5<EE>z^\_<A5><B7>A<94><AF>D<94><DE>^A<CA>n^<BE>l=^R<B0><97>

<D5>t<C4>^P<C3>[]<81><FE>[]<C0><F4><9F><D6><F9>50<9E><F0>J<B0><F8><ED><A7><C2>^M<B2>^E^A<AE>:i&^N^M^Ndvl^<BF>m<ED>#^L\$<DF>^L<D0>g<FC<A6>^HB\$z<C0>CME<dY^V<C1>eq<CA><CC>2<CC>^T<BF>,<)<EB><94>]<8C><ED>5\$f\$<A7><96><8C><D4>0RK<CB><C4>t<9A>[]<88><DD>^H^A^A2<A3><81><99>]<E3>=s<C7><FD>2<F2><CA>cESC^?<E3>^Gk<FF><CE>f<FD><8C><EF>]>^U^D<EC>^U<FF><CE>c<9E>ESC^G<EB><95><EB>^W<CF>x<9B><A7><U+0088><BD><E4><F6>^@<BD>uhw^r^LQJ<A7>,<C5>kCT<EB>br<C0><83>W<AD><AC>.<E0><89><E8>^T<95>!]<E3><97>T<98><E0>x^V<85><94>^T\*<D8>6r{<89><ED><AB><CB><E<CC><ED><C4>k<A8><FF>J<EB><AD>

<87><F0>^@B^@PS<A0><93><D0>^C<E6>[]o<9B><E4><B9>ESC<BC>KQ{<93>^^on<92><E7><E1>M<F2>[]<BD>K<8F><B7>lw<EE><83><E7><AD><DA>{( <F4>8^V:gp<CF><D7>[]^?^U<87><AB><C5><DB>^A<8E><99><9C>@<D0>^D<BA>8<B1><C4>\$MDI<CD>s3<C9>d<D1>^@<DB>^HN<9A><98>Tp^C<AA>DV<B3>J<AC>e9<AB><CD><CF><FA><B2><BD><EB><DC>[]<D4><DC>:I<84><FB><F5>U<A8><9F><BC>]]<B9>N<A2><8B><E8>d2^F<E7>#<B8><F4><C2>J<BC><DB>\x<9B>]pp<F1>[]<F0><EA><BC>^RSpk^-L^k<AB>L<E4>4+]<B2>RC2<E0>!"VY<9D>%<B5>QUI<A6>8<95>j^Bl^q.<3AC>^VZ^K[]<BC>4<92>\$<D5>U<AD>^S<CE>^^G<BD>siG[]<B2>zCk<FF>0^R<F0><C8>h<F1><80><EE>+<FF>^B<A3><FB>"]<A3>[]<DF>^

# FASTA

## >SEQUENCE\_1

TCGATACC-TGTCCAAAACAGAACGACCCGAGAACGATTGATCATCACTCTTGGCGGGGCC-  
GGTTTCTTAGCCGATTCTGTGCCCCGCTGATTCCGTTGTTTTGCTTATGGTTCCCGTCAGGT  
TCATTTTGAACATGATTGGACTATGCGCTTAGCTTTCGGAAATTCACAAAACCCCGGCACG  
AAAAGTGTCAAGGAACATGCAACTAAACAGCCTGCTTCCGCTGCCCCGGAGACGGTGAGT  
GTGCGGATGCTGTGCTGCGATCTAAAGTCTAAAACGACTCTCGGCAACGGATATCTCGGC  
TCTCGCATCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGTG  
AACCATCGAGTCTTTGAACGCAAGTTGCGCCCTAAGCCTTTTGGCCGAGGGGCACGTCTGC  
CTGGGTGTCACAAATCGTCGTCCCCAATCCTCTATGGATACAGGACGGAAACTGGTCTCC  
CGTGTGTTACCGTACGCGGTTGGCCAAAATCCGAGCTAAGGACGCCTGGAGTGTCTCGAC  
ATGCGGTGGTGAATTCAAGCCTCTTGATATTGTTGAACGCTCCTGTTCGAAGCTTTAGATG  
ACCCAAAGATATCAACGCGACCCCA

## >SEQUENCE\_2

TCGATACCTGTCCAAAACAGAACGACCCGAGAACGATTGATCATCACTCTCGGCGGGGCC-  
GGTTTCTTAGCCGATTCTGTGCCCCGCTGATTCCGTGGTTTTGCGAGTGGTTCTATGAGATT  
TTTAATCCTGATTGGGCTATGAGCTTAGCTTTCGGAAATTCACAAAACCCCGGCACGAAAA  
GTGTCAAGGAACATGCAACTAACAGCCTGCTTCCGCGCCCCCGGAAACGGTGAGTGTGC  
GGATGCTGTGCTGCGATCTAAAGTCTAAAACGACTCTCGGCAACGGATATCTCGGCTCTC  
GCATCGATGAAGAACGTAGCGAAATGCGATACTTGGTGTGAATTGCAGAATCCCGTGAAC  
CATCGAGTCTTTGAACGCAAGTTGCG-CCCTAAGCCTTCTGGCCGAGGGGCACGTCTGCC-  
TGGGTGTCACAAATCGTCGTCCCCAATCCTCTAAGGATAGAGGACGGAAACTGGTCTCCC  
GTGTGTTACCGTACGCGGTTGGCCAAAATCCGAGCTAAGGATGCCTGGAGCGTCTTTACA  
TGCGGTGGTGAATTCAAGCCTCTTGATATTGTTGAACGCTCCTGTCCGAAGCTTAAGATGA  
CCCAAAGACCTCAACGCGACCCCAGGTCAGGCGGGATCACCCGCTGAGTT

# Veri Tabanları

Veri Tabanı	Örnek
Bibliyografik VT	MEDLINE, Pubmed
Genom VT	Genome Information Broker (GIB), Entrez genome of NCBI
Sekans VT	DDBJ, EMBL, SWISS prot
Yapı (Structure) VT	Nucleotide Database (NDB), Protein Database (PDB)
Metabolik VT	Kyoto Encyclopedia of genes and genomes (KEGG)
Enzim VT	ExPasy, REBASE
Hastalık VT	OMIM
Kimyasal VT	PubChem
Microarray VT	Human Gene Expression Index (HUGE)

# Molecular Biology Databases

- **Nükleik asit dizileri**
- **GenBank** (1982, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)), **EMBL** (1982, [www.ebi.ac.uk](http://www.ebi.ac.uk)),
- **DDBJ** (1984, [www.ddbj.nig.ac.jp](http://www.ddbj.nig.ac.jp))
- **Amino asit dizileri:**
- **PIR** (1968, [www-nbrf.georgetown.edu](http://www-nbrf.georgetown.edu)), **PRF** (1979, [www.prf.op.jp](http://www.prf.op.jp)), **SWISS-PROT** (1986, [www.expasy.ch](http://www.expasy.ch))
- **3D Moleküler yapılar :**
- **PDB** (1971, [www.rcsb.org](http://www.rcsb.org)), **CSD** (1965, [www.ccdc.cam.ac.uk](http://www.ccdc.cam.ac.uk)) Lack standardization of data contents


# NCBI

ncbi.nlm.nih.gov/nuccore

Yandex Suggested Sites IE'den Al Molecular mar... PLANTS & FLO... Plant Myths a... UNIX / Linux T... panix UNIX/Linux In...

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Advanced Help



## Nucleotide

The Nucleotide database is a collection of sequences from several sources, including GenBank, RefSeq, TPA and PDB. Genome, gene and transcript sequence data provide the foundation for biomedical research and discovery.

### Using Nucleotide

- [Quick Start Guide](#)
- [FAQ](#)
- [Help](#)
- [GenBank FTP](#)
- [RefSeq FTP](#)

### Nucleotide Tools

- [Submit to GenBank](#)
- [LinkOut](#)
- [E-Utilities](#)
- [BLAST](#)
- [Batch Entrez](#)

### Other Resources

- [GenBank Home](#)
- [RefSeq Home](#)
- [Gene Home](#)
- [SRA Home](#)
- [INSDC](#)

You are here: NCBI > DNA & RNA > Nucleotide Database

Support Center

ki2013.pdf ^

# NCBI

→ [ncbi.nlm.nih.gov/nuccore/?term=Noccaea+papillosa](#) ☆

Apps Yandex Suggested Sites IE'den Al Molecular mar... PLANTS & FLO... Plant Myths a... UNIX / Linux T... UNIX/Linux In... » Other bookmark

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide   [Create alert](#) [Advanced](#) [Help](#)

Species  
Plants (4)  
[Customize ...](#)

Molecule types  
genomic DNA/RNA (4)  
[Customize ...](#)

Source databases  
INSDC (GenBank) (4)  
[Customize ...](#)

Sequence Type  
Nucleotide (4)

Genetic compartments  
Chloroplast (2)  
Plastid (2)

Sequence length  
[Custom range...](#)

Release date  
[Custom range...](#)

Revision date  
[Custom range...](#)

[Clear all](#)  
[Show additional filters](#)

Summary Sort by Default order Send to: [Manage Filters](#)

**Items: 4**

☐ [Noccaea papillosa voucher HUB:Ozudogru 3597 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene, complete sequence; and internal transcribed spacer 2, partial sequence](#)  
625 bp linear DNA  
Accession: MG944891.1 GI: 1603738095  
[Taxonomy](#)  
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

☐ [Noccaea papillosa voucher HUB:K.O 1083 tRNA-Leu \(trnL\) gene and trnL-trnF intergenic spacer, partial sequence; chloroplast](#)  
698 bp linear DNA  
Accession: MG925461.1 GI: 1558870779  
[Taxonomy](#)  
[GenBank](#) [FASTA](#) [Graphics](#) [PopSet](#)

☐ [Noccaea papillosa voucher HUB:Ozudogru 3597 tRNA-Leu \(trnL\) gene and trnL-trnF intergenic spacer, partial sequence; chloroplast](#)  
699 bp linear DNA  
Accession: MG925459.1 GI: 1558870777  
[Taxonomy](#)

**Analyze these sequences** Run BLAST

**Find related data** Database:  Find items

**Search details** Search See more...  
"Noccaea papillosa"  
[Organism] OR Noccaea papillosa[All Fields]

**Recent activity** Turn Off Clear



# NCBI

ncbi.nlm.nih.gov/nuccore/MG925461.1

Apps Yandex Suggested Sites IE'den Al Molecular mar... PLANTS & FLO... Plant Myths a... UNIX / Linux T... UNIX/Linux In... Other bookm

Nucleotide Nucleotide Search Advanced Help

GenBank ▾ Send to: ▾

**Noccaea papillosa voucher HUB:K.O 1083 tRNA-Leu (trnL) gene and trnL-trnF intergenic spacer, partial sequence; chloroplast**

GenBank: MG925461.1

**FASTA** Graphics PopSet

Go to: ▾

LOCUS MG925461 698 bp DNA linear PLN 20-DEC-2019

DEFINITION Noccaea papillosa voucher HUB:K.O 1083 tRNA-Leu (trnL) gene and trnL-trnF intergenic spacer, partial sequence; chloroplast.

ACCESSION MG925461

VERSION MG925461.1

KEYWORDS .

SOURCE chloroplast Noccaea papillosa

ORGANISM [Noccaea papillosa](#)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Gunneridae; Pentapetalae; rosids; malvids; Brassicales; Brassicaceae; Coluteocarpeae; Noccaea.

REFERENCE 1 (bases 1 to 698)

AUTHORS Ozudogru,B., Ozgisi,K., Tarikahya Hacioglu,B., Ocak,A., Lysak,M., Mummenhoff,K. and Al-Shehbaz,I.A.

TITLE Phylogeny of the genus Noccaea Moench (Brassicaceae) and a critical review of its generic circumscription

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 698)

Change region shown ▾

Customize view ▾

Analyze this sequence ▴

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

Related information ▴

Taxonomy

BioCollections

PopSet

Recent activity ▴

Turn Off Clear

Noccaea papillosa voucher HUB:K.O 1083 tRNA-Leu | Nucleotide



# NCBI

ncbi.nlm.nih.gov/nuccore/MG925461.1?report=fasta

Yandex Suggested Sites IE'den Al Molecular mar... PLANTS & FLO... Plant Myths a... UNIX/Linux T... UNIX/Linux In... Other bo

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Help

FASTA Send to: Change region shown Customize view

**Noccaea papillosa voucher HUB:K.O 1083 tRNA-Leu (trnL) gene and trnL-trnF intergenic spacer, partial sequence; chloroplast**

GenBank: MG925461.1

[GenBank](#) [Graphics](#) [PopSet](#)

```
>MG925461.1 Noccaea papillosa voucher HUB:K.O 1083 tRNA-Leu (trnL) gene and trnL-  
trnF intergenic spacer, partial sequence; chloroplast  
GAGCCTTGGTATGGAAACCTACTAAGTGATACTTTCAAATTCAGAGAAACCTGGAATTAACAATGGGC  
AATCCTGAGCCAAATCCTGGTTTACGCGCACAAACCGAGTTTAAAAAGCGAGAAAAGGGATAGGTGCA  
GAGACTCAATGGAAGCTGTTCTACAAATGGAGTTCACTACCTTGTTGATCAATGATTCACTTCATA  
GTCTGATAGATCCTTGGTGGAACTTATTAATCGGACGAGATAAAGATAGAGTCCATTCTACATGTCAA  
TACTGACAACAATGAAATTTATAGTAAGATGAAAATCCGTTGACTTTTTAAATCGTGAGGGTCAAGTCC  
CTCTATCCCCCACTCCCCGCGCACACCTTACCTTTTTTTAGTTATTCAAAAATTCATTATCTTTTTTCA  
TTCATCCTACGCTTTTACAACTAAAATTTCTTTCTTATTATACAAGTCTTGTGGGATATATCATCC  
ACGTACAAATGAGAAAGAAATATCGATTGAATTATTAGAATCTATATCATTTTCATTTCTAAAACCTT  
AGAAAGTCTTCTTTTCGCGAATCCAAGAAATCCCGGTCCAAAACCTTTTGCGTTTACTATTTTTTTTC  
GTTTCTTTTAATTGACATAGACCTAAGTCATCTAGTAAAATGAGGATGATACTTCGGTAATCGCCGGG
```

Analyze this sequence Run BLAST Pick Primers Highlight Sequence Features Find in this Sequence

Related information Taxonomy BioCollections PopSet

# Hesaplamalı Biyolojide kullanılan programlama dilleri

- Hangi programlama dilini öğrenmeliyim?

# Web uygulamaları



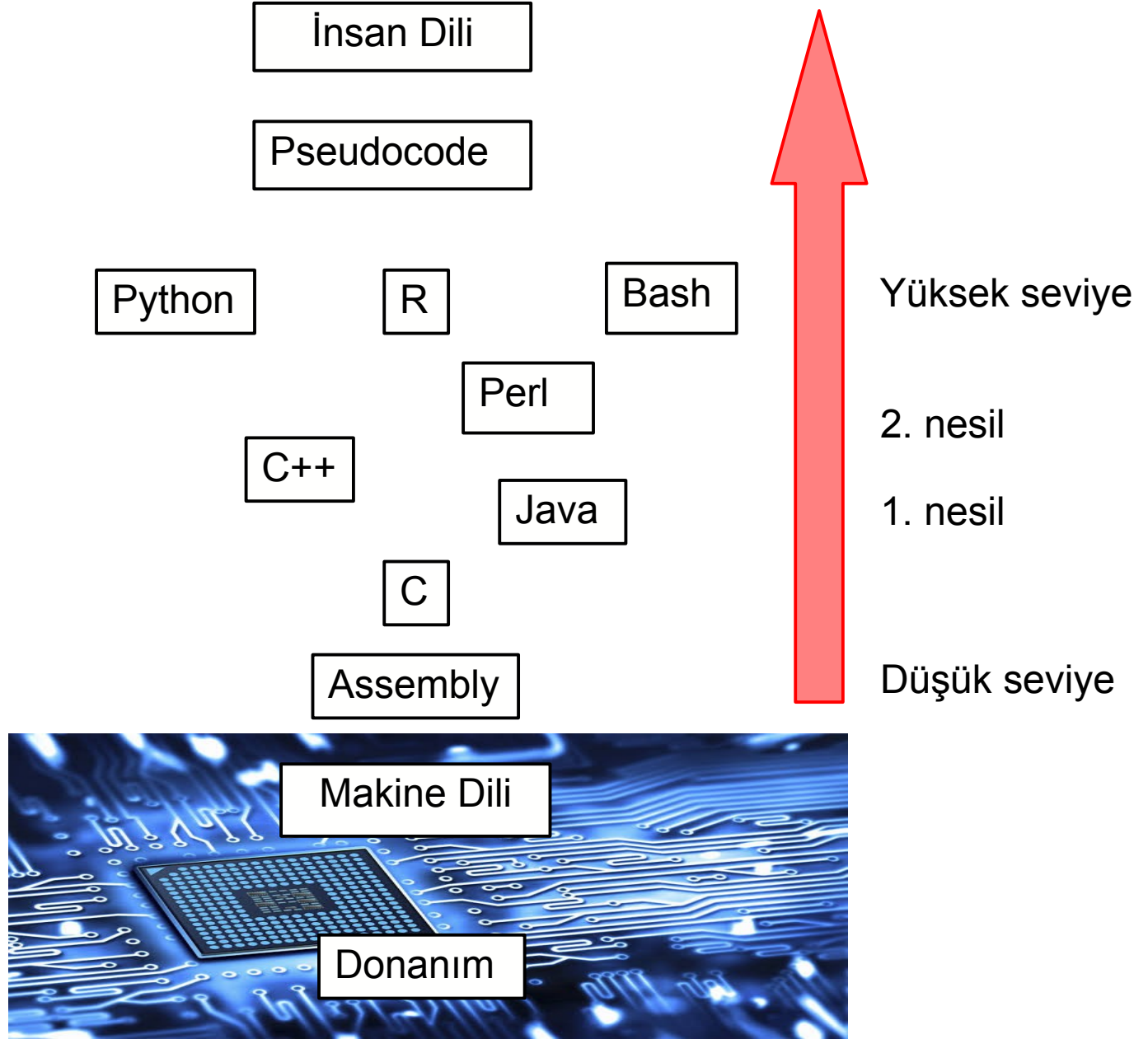
# Graphical User Interface (GUI)



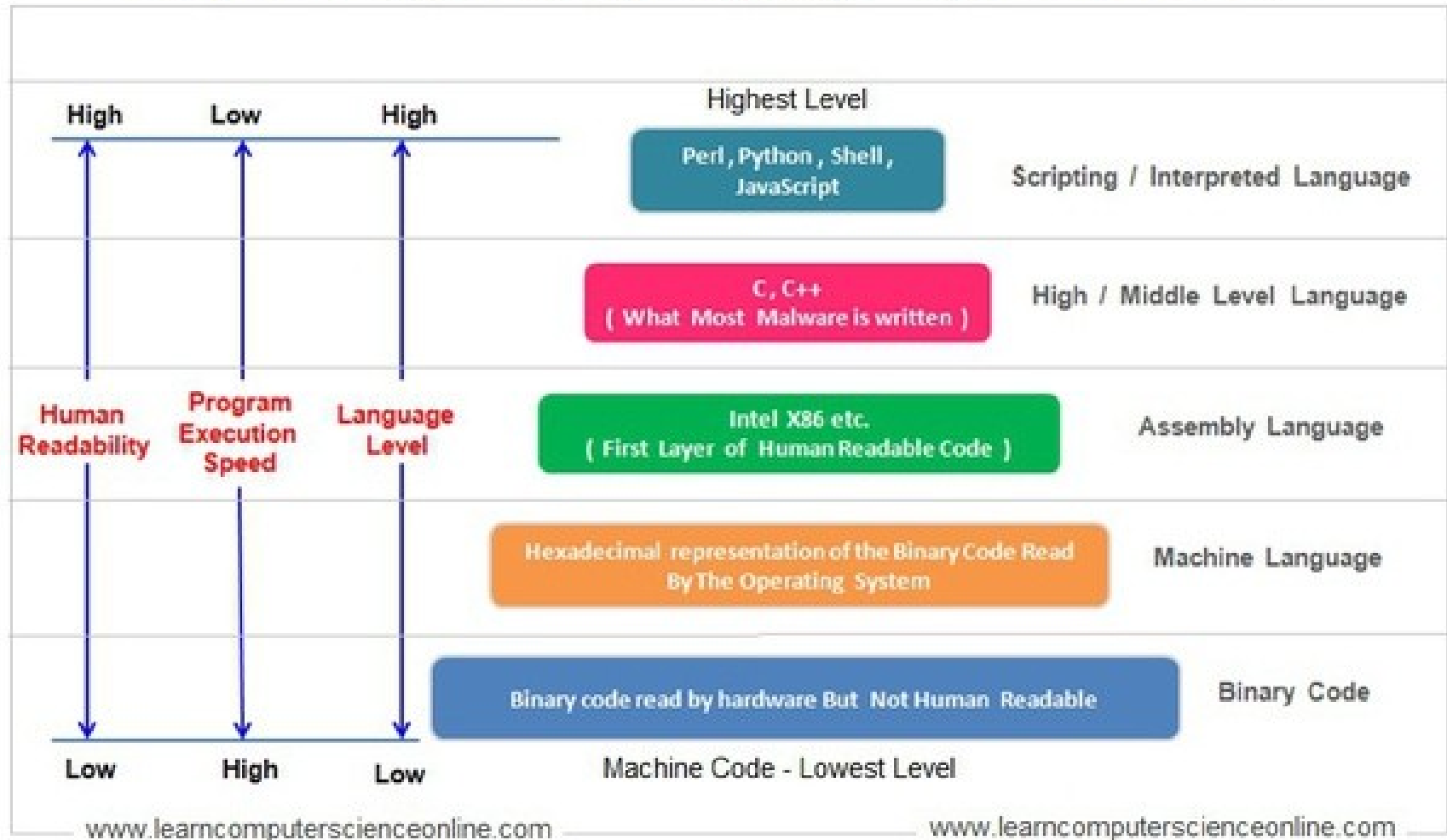
# Oyun geliştirme



# Programlama dilleri




# Computer Programming Language - Types And Levels



# Assembly

```
global _start
section .text
_start:
    mov rax, 1
    mov rdi, 1
    mov rsi, msg
    mov rdx, msglen
    syscall
    mov rax, 60
    mov rdi, 0
    syscall
section .rodata
msg: db "Hello, world!", 10
msglen: equ $ - msg
```

Everything Starts from a



<Hello World/>

@MUKUL

# C

```
#include <stdio.h>
int main()
{
    printf("Hello, world!");
    return 0;
}
```



# C++

```
#include <iostream>
int main()
{
    std::cout << "Hello, world!" << std::endl;
    return 0;
}
```





# Java

```
public class HelloWorld {  
    public static void main(String[] args) {  
        System.out.println("Hello, world");  
    }  
}
```



# Python, R, Bash

```
print("Hello, world!")
```



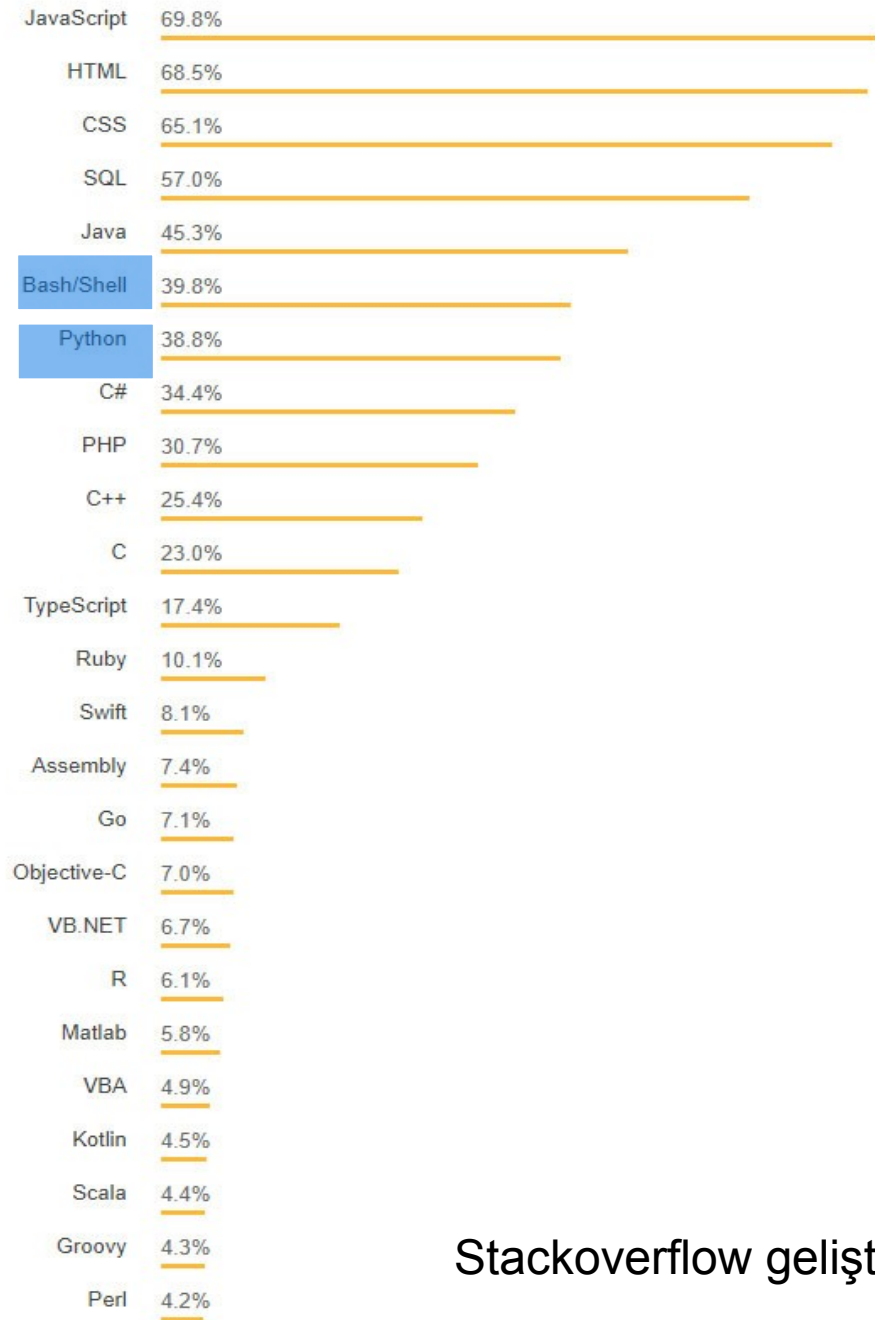
```
print("Hello, world!")
```



```
echo "Hello, world!"
```



```
#!/bin/bash
```



Stackoverflow geliştirici anketi (2019) sonuçları

Jan 2020	Jan 2019	Change	Programming Language	Ratings	Change
1	1		Java	16.896%	-0.01%
2	2		C	15.773%	+2.44%
3	3		Python	9.704%	+1.41%
4	4		C++	5.574%	-2.58%
5	7	^	C#	5.349%	+2.07%
6	5	v	Visual Basic .NET	5.287%	-1.17%
7	6	v	JavaScript	2.451%	-0.85%
8	8		PHP	2.405%	-0.28%
9	15	^^	Swift	1.795%	+0.61%
10	9	v	SQL	1.504%	-0.77%
11	18	^^	Ruby	1.063%	-0.03%
12	17	^^	Delphi/Object Pascal	0.997%	-0.10%
13	10	v	Objective-C	0.929%	-0.85%
14	16	^	Go	0.900%	-0.22%
15	14	v	Assembly language	0.877%	-0.32%
16	20	^^	Visual Basic	0.831%	-0.20%
17	25	^^	D	0.825%	+0.25%
18	12	vv	R	0.808%	-0.52%
19	13	vv	Perl	0.746%	-0.48%
20	11	vv	MATLAB	0.737%	-0.76%

TIOBE Programlama Topluluğu indeksine göre en yaygın kullanılan programlama dilleri

# Bazı programlama dilleri ve kullanım alanları



Python

## Kullanım alanları:

- Web geliştirmeleri
- Video oyun geliştirme
- Yazılım geliştirme
- Veritabanı erişimi

## Popüler uygulamalar

➤ Instangram



➤ YouTube



➤ Spotify



➤ Reddit



# Bazı programlama dilleri ve kullanım alanları



## Kullanım alanları:

- Android & IOS uygulama geliştirmeleri
- Video oyun geliştirme
- Yazılım geliştirme

## Popüler uygulamalar

- G-Mail



- Minecraft



- Android



# Bazı programlama dilleri ve kullanım alanları



## Kullanım alanları:

- Yazılım geliştirmeleri
- Video oyun geliştirme
- İşletim sistemi yazılımı
- Arama motoru geliştirme

## Popüler uygulamalar

- Google



- Photoshop

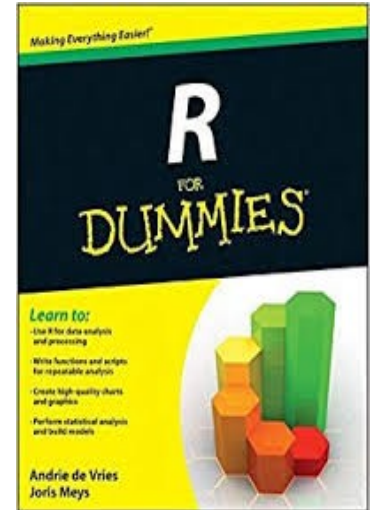
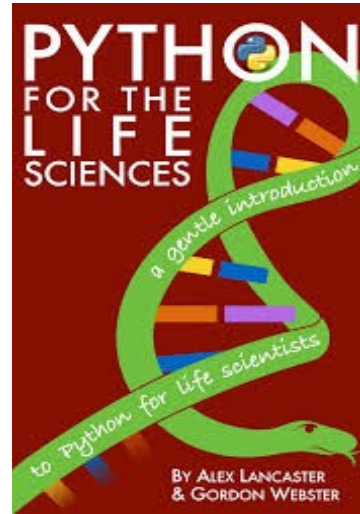
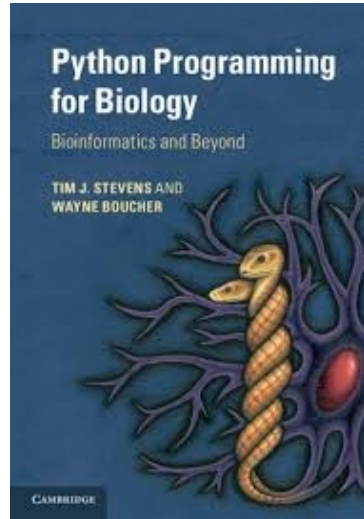


- Mozilla Firefox



- Yahoo





Temel okuma önerileri



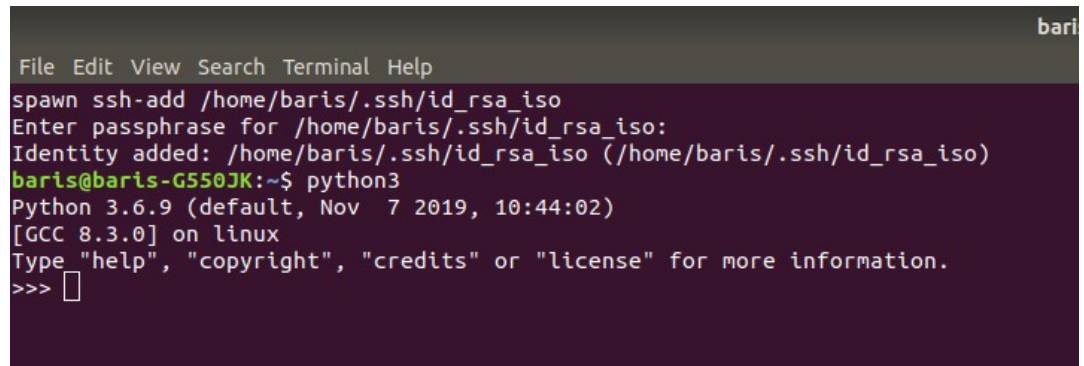
# Neden Python?

- Python, öğrenmesi kolay, tamamen özgür ve ücretsiz bir programlama dilidir.
- Nesne yönelimli bir dil olup okunabilirliği yüksektir.
- Birçok programlama diline göre daha az kod ile işlemler yapılabilir.
- Python, bütün işletim sistemleri ile uyum içerisinde çalışmaktadır.
- Programlama yapısı içerisinde birçok kütüphaneyi barındırmaktadır.
- Python ile masaüstünde çalışan uygulamalar geliştirilebileceği gibi, web üzerinde çalışan uygulamalar geliştirmek hatta Raspberry-Pi gibi donanımları da programlamak mümkündür.

# Python programlama diline giriş

Mac ve Linux için:

/usr/local/bin/python3 ya da

A terminal window with a dark purple background and white text. The window title is 'bari'. The menu bar at the top includes 'File', 'Edit', 'View', 'Search', 'Terminal', and 'Help'. The terminal output shows the command 'ssh-add /home/baris/.ssh/id\_rsa\_iso' being executed, followed by a prompt for a passphrase. After the passphrase is entered, the message 'Identity added: /home/baris/.ssh/id\_rsa\_iso (/home/baris/.ssh/id\_rsa\_iso)' is displayed. Then, the command 'python3' is entered, resulting in the output 'Python 3.6.9 (default, Nov 7 2019, 10:44:02) [GCC 8.3.0] on linux'. A final prompt asks the user to type 'help', 'copyright', 'credits', or 'license' for more information. The prompt '>>>' is followed by a cursor icon.

```
File Edit View Search Terminal Help
spawn ssh-add /home/baris/.ssh/id_rsa_iso
Enter passphrase for /home/baris/.ssh/id_rsa_iso:
Identity added: /home/baris/.ssh/id_rsa_iso (/home/baris/.ssh/id_rsa_iso)
baris@baris-G550JK:~$ python3
Python 3.6.9 (default, Nov 7 2019, 10:44:02)
[GCC 8.3.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> █
```

# Terminoloji

**String (karakter dizileri):** Sabit veya değişken olabilen bir veya daha fazla karakterden (harfler, sayılar, simgeler) oluşan bir dizidir.

```
>>> s1 = "string",  
>>> s2="123",  
>>> s3 ="hesaplamalı biyoloji"  
>>> print(s1, s2, s3)  
string 123 hesaplamalı biyoloji
```

**Değişken:** Değerleri depolamak için ayrılmış bellek konumlarıdır.

```
>>> var =30  
>>> print(var)  
30
```

```
>>> var = "şimdi de değişken bir karakter dizisi olsun"  
>>> print(var)  
şimdi de değişken bir karakter dizisi olsun
```

# Terminoloji

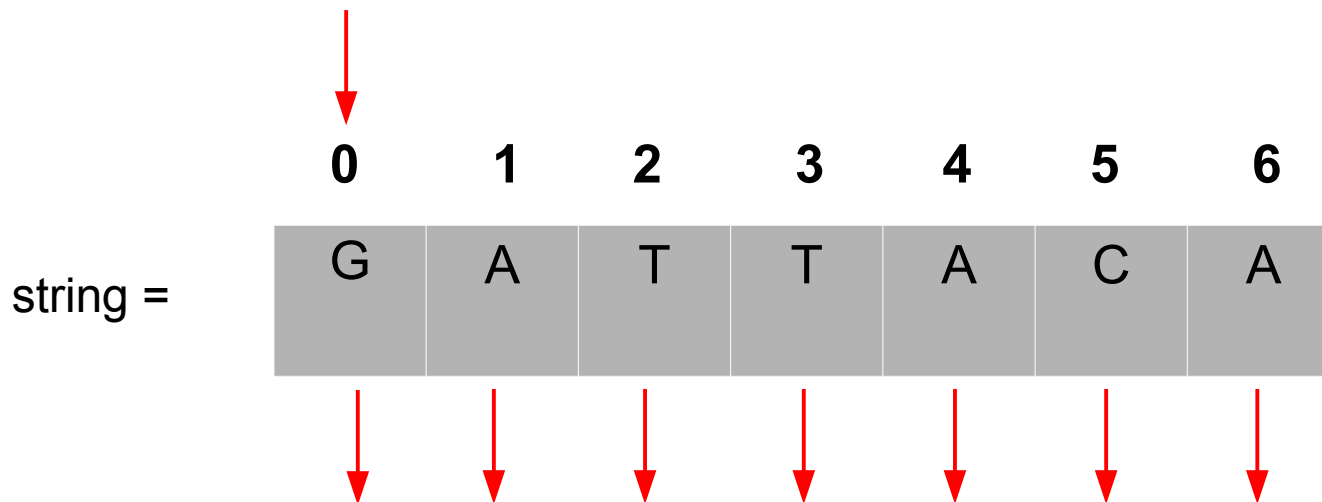
**Fonksiyon:** Girdi alan, belirli hesaplamalar yapan ve çıktı üreten bir ifade kümesidir.

```
>>> def ekrana_yaz(str):  
    print(str)  
    return  
ekrana_yaz(Merhaba")
```

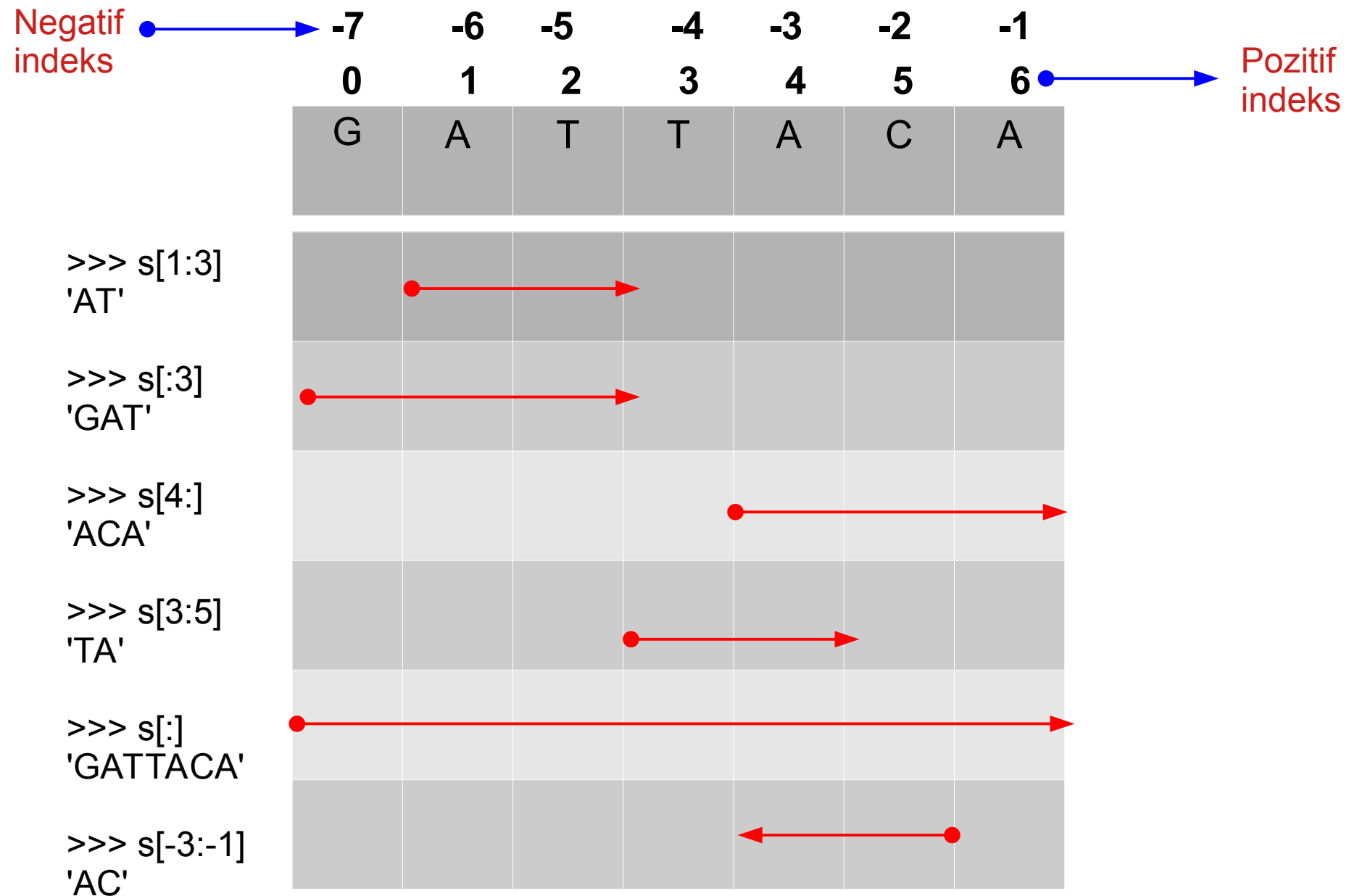
# String

```
>>> s = "GATTACA"
```

**Önemli!!!**



# İndeksler



# Neden stringler önemlidir?

→ DNA dizileri string'tir.

catgaaggaa ccacagccca gagcaccaag

→ Veri tabanı kayıtları string'lerden oluşur.

LOCUS JX915225 889 bp DNA linear PLN 12-DEC-2012

DEFINITION *Campanula alliariifolia* isolate NC3383 petB-petD intergenic  
spacer,partial sequence; and petD protein (petD) gene, exon  
and partial cds; chloroplast.

ACCESSION JX915225

VERSION JX915225.1

.....

→ HTML büyük bir string'tir.

# String'leri oluşturma

“bu bir string”

“bu başka bir string”

“stringler çift tırnak işaretleri arasında bulunur”

‘ya da tek tırnak işaretleri arasında bulunabilir’

‘bunların arasında bir fark yoktur’

‘Okay, there\'s a small one.’



# String'in bir kısmını ayıklama (çıkarma =extracting)

```
>>>protein = "vlspadktnv"
```

# 3'ten 5' e kadar olan pozisyonları yazdıralım

```
>>>print(protein[3:5])
```

pa



?

# pozisyonlar 0 dan itibaren başlar ve başlangıçta kapsayıcı bitişte hariç bırakıcıdır.

```
>>>print(protein[0:6])
```

vlspad

# kısa bir DNA dizisini **my\_dna** değişkeni içerisinde saklayalım

```
>>>my_dna = "ATGCGTA"
```

# şimdi DNA dizisini yazdıralım

```
>>>print(my_dna)  
ATGCGTA
```

Değişken ismi tamamen keyfidir. İsteddiğiniz herhangi bir ismi verebilirsiniz.

```
>>>jon_snow= "ATGCGTA"  
>>>print(jon_snow)  
ATGCGTA
```

! Değişken isimleri BÜYÜK - küçük harfe duyarlıdır.

Yani;

my\_dna,

MY\_DNA,

My\_DNA,

My\_Dna farklı değişkenlerdir.

# Stringleri birleştirme

```
>>>my_dna = "AATT" + "GGCC"  
>>>print(my_dna)  
AATTGGCC
```

Bu örnekte iki string birleştirilmiştir, ancak stringleri işaret eden değişkenler de birleştirilebilir.

```
>>>upstream = "AAA"  
>>>my_dna = upstream + "ATGC"  
>>>print(my_dna)  
AAAATGC
```

# Birkaç örnek

```
>>> len("GATTACA")
7
>>> "GAT" + "TACA"
'GATTACA'
>>> "A" * 10
'AAAAAAAAAAAA'
>>> "G" in "GATTACA"
True
>>> "GAT" in "GATTACA"
True
>>> "AGT" in "GATTACA"
False
>>> "GATTACA".find("ATT")
1
>>> "GATTACA".count("T")
2
```

# Büyük-küçük harf değiştirme

## *lower()* ve *upper()* fonksiyonları

```
>>>my_dna = "ATGC"  
>>>print(my_dna.lower())  
atcg
```

Ya da

```
>>>my_dna = "atgc"  
>>>print(my_dna.upper())  
ATGC
```

# Yenisıyla değiştirme

## *replace* fonksiyonu

```
>>>protein = "vlspadktnv"
```

# şimdi valin'i (v) tirozinle (y) değiştirelim

```
>>>print(protein.replace("v", "y"))  
ylspadktny
```

# burada birden fazla karakteri de değiştirebiliriz.

```
>>>print(protein.replace("vls", "ymt"))  
ymtpadktnv
```

# substring'leri sayma ve bulma

```
>>>protein = "vlspadktnv"
```

# burada istediğimiz bir amino asidi saydırabiliriz

```
>>> valine_count=protein.count('v')  
>>>print(valine_count)  
2
```

```
>>>lsp_count = protein.count('lsp')  
  
>>>print(lsp_count)  
1
```

```
>>>tryptophan_count = protein.count('w')  
  
>>>print(tryptophan_count)  
0
```



# Dosyadan metinleri okuma

## *open ve read* fonksiyonları

Metin dosyaları nelerdir?

- DNA ya da protein dizilerine ait FASTA dosyaları
- komut tabanlı program çıktıları (e.g. BLAST)
- DNA okumaları içeren FASTQ dosyaları
- HTML dosyaları
- Python kodları vs.

```
>>>my_file = open("dna.txt")
>>>file_contents=my_file.read()
>>>print(file_contents)
```

# Dosyadan metinleri okuma

## *open* ve *read* fonksiyonları

```
1 >>>my_file_name = "dna.txt"
2 >>>my_file = open(my_file_name)
3 >>>my_file_contents = my_file.read()
  ACTGTACGTGCACTGATC
```

# 1. satırda dna.txt stringi **my\_file\_name** değişkeni içerisinde saklanıyor.

# 2. satırda my\_file\_name değişkenini **open** fonksiyonu için bir argüman olarak kullanıyor ve sonucu **my\_file** değişkeni içerisinde saklıyoruz.

# 3. satırda **my\_file** değişkenindeki **read** metodunu çağırıyoruz ve elde edilen stringi **my\_file\_contents** değişkeninde saklıyoruz.