



<http://lectures.molgen.mpg.de/statistik/>

Ein Vergleich von Shakespeare, Goethe und dem menschlichen Genom



Genomische Datenanalyse

1. Kapitel

Genomische Daten

ttgcgccgccaacgcccggactggagccctgacattgctgcgctcggggggctcca
 ggcagcccgtatcggggccttatttctcgtcggcggcgcctgcaccgcccctcgaaa
 gccaaacaccagagcaccctagaaggtttaaactaaaagaaagtctcatggttgaccagtt
 cctgtcaagcaagaggccaaggaccctgctcagtgctataccacttaattacatggaa
 tccaatgaagcctaacaagtaggggtcatctactccacaccattgcctgagaagttcttt
 cagacccagaaggtctgtcgcacggaaacagatggagccagtgaccctcacggtgaac
 aagcggagttcaccctctcggctgggaaatcgccctcctctctgaagttcccgtcctca
 caccggagagcctcggctgggttgagcatgacctcttccagcccaccgataaaaaaac
 tcaccctctctccaggcgtgcagccctcggcgtgcgcgtgtccaagccaccagtgatg
 gcagctgccctctcgcggcaaggaaacggagcccggggatactgcccgtcatccagccg
 gtggtggtgcagcccgtccccttattgtacacaagtcacctccagcagcctctcaatggtc
 tccttatcggaggagatggaaaaatccagtagtagcatgcaagtcacctgtaattgaaatca
 tatgagaagcctatatcacagaaaaaaataaaaatagaacctgggatacgaaccacagag

Erster Eindruck

Das ist eine völlig unverständliche, scheinbar zufällige Anordnung von a,c,g,t

ttgcgccgccaacgcccggactggagccctgacattgctgcgctcggggggctcca
 ggcagcccgtatcggggccttatttctcgtcggcggcgcctgcaccgcccctcgaaa
 gccaaacaccagagcaccctagaaggtttaaactaaaagaaagtctcatggttgaccagtt
 cctgtcaagcaagaggccaaggaccctgctcagtgctataccacttaattacatggaa
 tccaatgaagcctaacaagtaggggtcatctactccacaccattgcctgagaagttcttt
 cagacccagaaggtctgtcgcacggaaacagatggagccagtgaccctcacggtgaac
 aagcggagttcaccctctcggctgggaaatcgccctcctctctgaagttcccgtcctca
 caccggagagcctcggctgggttgagcatgacctcttccagcccaccgataaaaaaac
 tcaccctctctccaggcgtgcagccctcggcgtgcgcgtgtccaagccaccagtgatg
 gcagctgccctctcgcggcaaggaaacggagcccggggatactgcccgtcatccagccg
 gtggtggtgcagcccgtccccttattgtacacaagtcacctccagcagcctctcaatggtc
 tccttatcggaggagatggaaaaatccagtagtagcatgcaagtcacctgtaattgaaatca
 tatgagaagcctatatcacagaaaaaaataaaaatagaacctgggatacgaaccacagag

„Die DNA ist das Buch des Lebens“

Die DNA codiert:

- Bauanleitungen für Proteine
- wann Proteine gebaut werden
- wie eine Zelle organisiert ist
- wie ein neuer Organismus entsteht

Wenn DNA codierend ist...dann sind die a,g,c,t **auf keinen Fall zufällig!**

Es gibt:

- Bedeutungen
- Struktur
- Grammatik(?)

Wo ist hier die Struktur?

```
ttgcgccgccaacgcgcccgactggagccctgacatgctgctcggggggtcca  
ggcagcccgtatcggggcttatttctcgtcggcgccctgcaccgcctcgcaa  
gccaaacaccagagcacccctagaaggtttaactaaaagaaatgctcatgtttgacccagtt  
cctgtcaagcaagaggccaaggaccctgtctcagtgctacaccaatctaatatacaggaa  
tccatgaagcctaacaagataggggtcactctactccacacattgctgagaagttcttt  
cagaccccagaaggtctgtcgcacggaatacagatggagccagtgacctcacggtgaa  
aagcggagttcaccctctcggctgggaattcgccctcctctctgaagttcccgctcctca  
caccggagagcctcgcctgggttgagcatgcctctctccagcccaccgataaaaaatac  
tcaccctctctccaggcgtgcagccctcggcgtgcccgtgtccatgccaccagtgatg  
gcagctgcccctctcggcagtggaatacggagcccggggatcctgccctcatccagccg  
gtggtggtgcagcccctccctttatgtacacaagtcacctccagcagcctctcatggtc  
tccttatcggaggagatggaaaattccagtagtagcatgcaagtacctgtaattgaaatca  
tatgagaagcctatatcacagaaaaaattaaaatagaacctgggatcgaaaccagagg
```

Wie sucht man nach einer Struktur?

Welches ist der strukturierte Text?

ihrnautechwiederschwankendegestaltendiefrüh
sich ein st dem trüben blick zeigt versuch ich wohl
euch dies mal fest zu halten fühlich mein herz noch je
nem wahngeneigt ihr drängt euch zu ungunst somögt ihr
walten wie ihr aus dunst und nebel um mich steigt mein
busen fühlt sich jugendlicher schüttert vom zauber
hauch der euren zugum wittert

tzeksbtstfdppoejjstqorjlstgpuettewyxxlüppt
ffhzdfhfuiddkdtwnnsbpawxvhszejketgfsdpovdk
uswkldjftehssjdgftrbxlficgeoszksezfjdbskxz
skdzzgsovkjshdxjxdjcuynzrksssdudjjskdasyd
hhzwli snfhzskqanxijndeidnrzbtflkpeqdoetakbr
klkderjsoendfuwtqkdhfdtwerrwqdhertqrlldjtanue
oethwieprnsiieserwqmqkspezu

ihrnautechwiederschwankendegestaltendiefrüh
sich ein st dem trüben blick zeigt versuch ich wohl
euch dies mal fest zu halten fühlich mein herz noch je
nem wahngeneigt ihr drängt euch zu ungunst somögt ihr
walten wie ihr aus dunst und nebel um mich steigt mein
busen fühlt sich jugendlicher schüttert vom zauber
hauch der euren zugum wittert

Wir können einen Text verstehen und erkennen
dadurch die Strukturen der Buchstabensequenz

Was hilft uns das bei der DNA?

Wir verstehen ja noch kein „genomisch“

Wie sieht es hier aus?

hikayatpandawainisangatterkenalterbukti
denganbanyaknyanaskahhikayatiniseperth
ikayatpandawalimahikayatpandawajayapand
awapancakelima.

uswkldjftehssjdgftrbxlficgeoszqksezfjbskx
zskdzzgsovkjshdzxsxdjcuynzrksssdudjjskdas
ydhzwlisnfhzskqanxijdndeidnrzrtbflkpeqdoeta
kbrklkdterjsoendfuwtqkdhfdtwerrwqdhqrldjt
anue

Welche Sequenz sieht strukturierter
und welche zufälliger aus?

Was macht hier Struktur aus?

Häufige Buchstaben

hikayatpandawainisangatterkenalterbukti
denganbanyaknyanaskahhikayatiniseperth
ikayatpandawalimahikayatpandawajayapand
awapancakelima.

Vermiedene Buchstabenfolgen

jktdffrhzrfgtewqq...

Wiederholungen

hikayatpandawainisangatterkenalterbukti
denganbanyaknyanaskahhikayatiniseperth
ikayatpandawalimahikayatpandawajayapand
awapancakelima.

Häufige Buchstaben

Vermiedene Buchstabenfolgen

Wiederholungen

Das sind Strukturmerkmale im Malaiischen.

Gilt das denn auch fürs Genomische?

Ja...mehr dazu später.

Zunächst kommen wir noch zu:



Was ist von Goethe und was von Shakespeare?

Dämmungsenktesichvonoben,
SchonistalleNähefern;
Dochzuerstemporgehoben
HoldenLichtsderAbendstern!
AllesschwanktinsUngewisse,
NebelschleichenindieHöh;
SchwarzvertiefteFinsternisse
WiderspiegelndruhtderSee.
NunimöstlichenBereiche
AhdichMondenglanzund-glut,
SchlankerWeidenHaargezweige
ScherzenaufdernächstenFlut.
DurchbewegterSchattenSpiele
ZittertLunasZauberschein
UnddurchsAugeschleichtdieKühle
SänftigendinsHerzhinein.

Mymistress'eyesarenothinglikethesun;
Coralisfarmoredthanherlips'red;
Ifsnowbewhite,whythenherbreastsaredun;
Ifhairsbewires,blackwiresgrowonherhead.
Ihaveseenrosesdamask'd,redandwhite,
Butnosuchrosesseelinercheeks;
Andinsomeperfumesistheremoredelight
Thaninthebreaththatfrommymistressreeks.
Ilovetohearherspeak,yetwellIknow
Thatmusicathafarmorepleasingsound;
IgrantIneversawagoddessgo;
Mymistress,whenshewalks,treadsontheground:
Andyet,byheaven,Ithinkmyloveasrare
Asanyshebeliedwithfalsecompare.

Ok, das ist zu einfach. Aber wie würden Sie die
Unterscheidung einem Computer beibringen?

Wie schreibt man ein Computerprogramm das Shakespeare Sonette von Goethe Gedichten unterscheiden kann?

Ansatz:

Goethe schreibt in Deutsch

Shakespeare in Englisch

Es gibt einfache charakteristische Unterschiede dieser Sprachen, die auch ein Computer finden kann

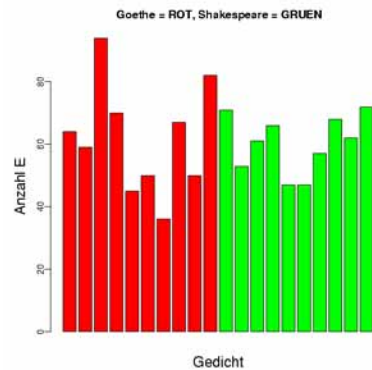
1. Schritt: Wir erheben Daten

- 10 Goethe Gedichte
- 10 Shakespeare Sonette

Dämmerung senkte sich von oben	Sonnet 130
Hab ich tausendmal geschworen	Sonnet 131
Zigeunerlied	Sonnet 132
Heidenröslein	Sonnet 133
Ein großer Teich war zugefroren	Sonnet 134
Fliegentod	Sonnet 135
Gefunden	Sonnet 136
Mignon	Sonnet 137
Nun sitzt der Ritter an dem Ort	Sonnet 138
Es war einmal ein König	Sonnet 139

Versuchen wir es damit einzelne Buchstaben zu zählen...

...zunächst den Buchstaben E:



Für jedes Gedicht wurde das Vorkommen des Buchstabens „E“ gezählt

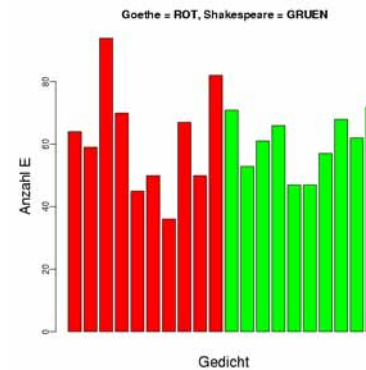
Dann wurden diese 20 Zahlen in einem Balkendiagramm dargestellt

Jeder Balken entspricht einem Gedicht

Die Höhe des Balken entspricht der Anzahl E's in diesem Gedicht

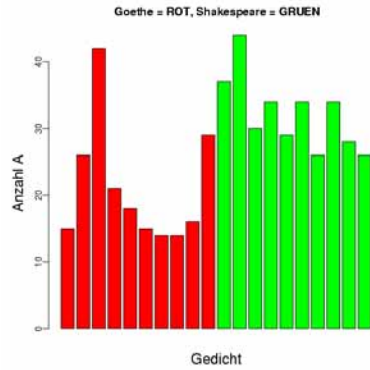
Goethe Gedichte sind durch rote Balken dargestellt, Shakespeare Sonetten durch grüne Balken

Beobachtungen



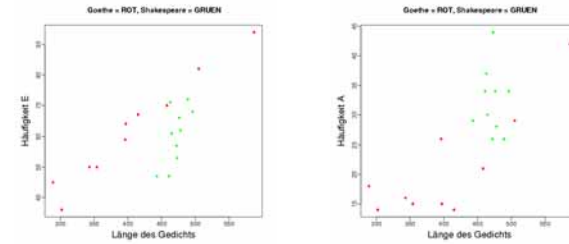
- Die meisten E's hat ein Goethe Gedicht
- Die wenigsten E's hat ebenfalls ein Goethe Gedicht
- Man kann nicht sagen, daß Goethe Gedichte mehr E's haben als Shakespeare Sonette oder umgekehrt
- Die Anzahl E's scheint in den Goethe Gedichten stärker zu streuen ?

Versuchen wir es jetzt mit den A's



- Goethe scheint **im Mittel** weniger A's zu benutzen als Shakespeare
- Aber nicht immer
- Wieder scheint die **Streuung** höher bei Goethe
- Woran kann das liegen?

Die Anzahl von A's und E's **korrelieren** mit der Länge des Gedichts



Im **Scatterplot** werden zwei Größen gegeneinander abgetragen

Jeder Punkt entspricht einem Gedicht

Der x-Achsenwert entspricht der Länge des Gedichts, der y-Achsenwert der Anzahl von E's in diesem Gedicht

Die Goethe Gedichte haben sehr unterschiedliche Längen

Die Längen von Shakespeare Sonetten streuen weniger (Def. Sonett)

In langen Goethe Gedichten sind sowohl mehr A's als auch E's als in kurzen (kein Wunder)

Für unser Computerprogramm ist das ein nicht gewünschter Effekt

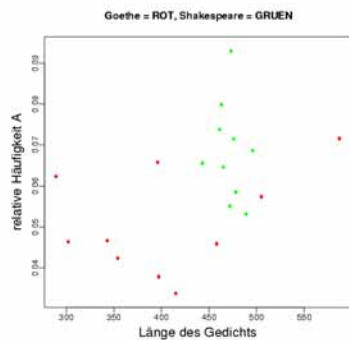
Die **absoluten Häufigkeiten** von A's und E's in den Gedichten sind nicht zu vergleichen

Berechne **relative Häufigkeiten** (zunächst für A)

relative Häufigkeit eines A=Anzahl A's/Gesamtanzahl Buchstaben

oder teilen durch die Länge

die relative Häufigkeit ist immer eine Zahl zwischen 0 und 1



Hier ist immer noch eine Tendenz zu erkennen, daß längere Gedichte auch relativ mehr A's enthalten

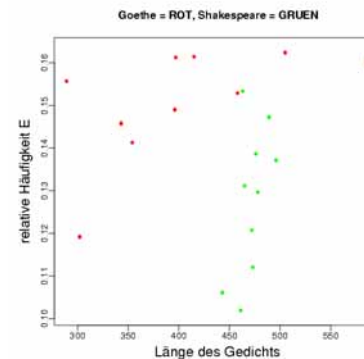
Aber nur wenn man die Goethe und die Shakespeare Gedichte gleichzeitig anschaut

Innerhalb der Gruppen sieht man keine Tendenz

Die relativen Häufigkeiten sind vergleichbar

Was man sieht ist, daß die betrachteten Shakespeare Sonette im Mittel länger sind als die Goethe Gedichte und auch im Mittel relativ mehr A's enthalten

Berechne **relative Häufigkeiten** (jetzt für E's)

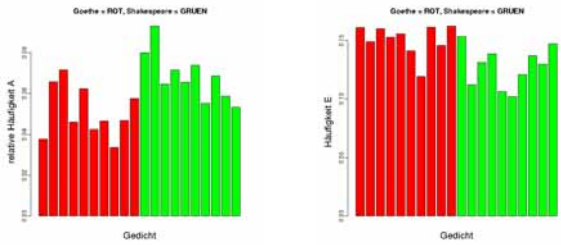


Hier scheint man die umgekehrte Tendenz zu erkennen, daß längere Gedichte relativ weniger E's enthalten

Aber auch das ist nur eine **Scheinkorrelation**, denn innerhalb der Farbgruppen existiert dieser Effekt nicht

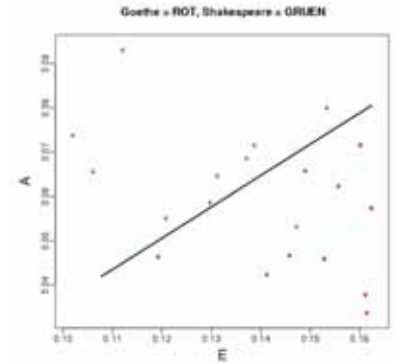
Was man sieht ist wiederum, daß die betrachteten Shakespeare Sonette im Mittel länger sind als die Goethe Gedichte, aber relativ weniger E's enthalten

Barplots der relativen A und E Häufigkeiten...



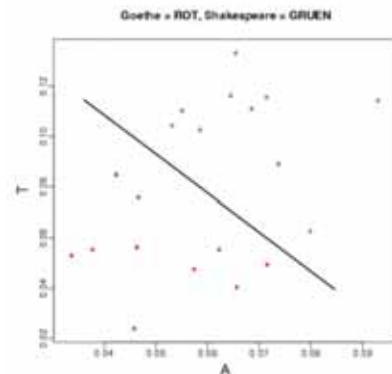
...zeigen, daß beide relativen Häufigkeiten nicht ausreichen um Goethe von Shakespeare zu unterscheiden...

...aber ein Scatterplot beider relativen Häufigkeiten gleichzeitig würde nur ein Shakespeare Sonett Goethe zuschreiben...



Mann kann die beiden Gruppen mit einer Trenngerade unterscheiden

...noch besser funktioniert es mit A und T



Offene Frage: Wie bringt man einem Computer bei, diese Trenngeraden zu berechnen?

Antwort: In Statistik 2

Was hat das alles mit Genomen zu tun?



Zwei Abschnitte aus dem menschlichen Genom

A

```
gtgcctcctgattaagagaaagatgtatattccactg
ggggcaggagacatagatgtccaggcccatgtgtccac
ggcaataatgtgggactgcagagaggaaggacatgttt
atgtcccagggtgagtcattggaatctcagaaatgagggt
gtgcatgacagatttttccaaggagggtgtcttagtca
cttcaggctgctaatacaaaatgccataaactgggtac
cttaaacacaacacttactcctcacagtgtcggaggc
aggaagtcatcaaggagctggaacattggagtctg
ccaaggcccacttctgtgtcttagccatcttctgtct
gtgttctcacctgggtggaaggatggggctgtctct
ggggtttcaaaaggcactaaatcccttcatgagggttc
cgtcctctgacctaaacccccaaaggccccacctct
aaataccatcacattgggggttagatttcaacataggaa
ttttgagagatgtgaacattttctgtgtgctgtaag
gggcatctcagagtgcacccccacatctcacaagctgc
tggaccttgcataatgggagctgactctcttgggctgtg
gtgactcatggctggagcctggctgttaaccagctc
tgaggctctctgggtgccacctctgcacaccagccagc
caggtggcctccacgtaccttctgtttcatggagaag
ctactgtcttaagacaggttctcaaacctggcatct
ggctcagggtgactctgagcatctcatggctacctcctc
cccactgcagcctcatctccaccctgcgcactctatg
actgagcagcacagggcgcca
```

B

```
cggtcactcgtgagcagtttct
gagtgacttctccccttctctggg
cgtcaaggcgaagcgtccgcgtta
gaaagaggaactccagttttaccga
agacctcaaggctctgcaaggagat
aacctcccggggagccatgcgcc
cgggtccagcggcctcccagcccgc
ggagcctcaaacctcgcggggccg
gaagcccggcggggaagcgggtg
tgcttttacgtccgccccgcgca
gccgcgaccgctgcggcgtctccg
cctgcctcctcgcggcggctcca
gtgcccgtctagagggcgtcctg
ggctagcgtgtagggctggcggcgg
cggcgtcgggtcacctctgggagcg
gagtgccggcggagcagagcgaag
cagctcaggagacttgaggcgtagg
ctccgctcccaagggtaccgcgcc
ctaigtgggactcgcctca
```

Informationen zur Sequenz A (Exon) aus der EMBL Datenbank

<http://srs.embl-heidelberg.de:8000/srs5/>

```
....
ID HS07000 standard; DNA; HUM; 152141 BP.
AC U07000;
SV U07000.1
DT 02-JUN-1994 (Rel. 39, Created)
DT 29-MAY-2002 (Rel. 71, Last updated, Version 37)
DE Human breakpoint cluster region (BCR) gene, complete cds.
OS Homo sapiens (human)
....
RX MEDLINE; 95394474.
RA Chisoe S.L., Bodenteich A., Wang Y., Jian L., Burian D., Carbtree J.S.,
RA Freeman A., Hamed O.S., McLaurie H., Pan H., Toth S., Wang Z., Zhang G.,
RA Clifton S., Groffen J., Heisterkamp N., Roe B.A.;
RT "Sequence and analysis of the human ABL gene, the BCR gene, and regions
RT involved in the Philadelphia chromosomal translocation";
RL Genomics 27(1):67-82(1995).
....
FT source 1..152141
FT /chromosome="22"
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /map="22q11,
...
FT conflict 32500
FT /replace="g"
FT /citation=[9]
FT /gene="BCR"
FT repeat_region 32914..33214
FT /citation=[29]
FT /rpt_family="Alu"
FT exon 33289..33430
FT /note="alternative exon 2"
FT /number=2
FT /gene="BCR"
....
```

Informationen zur Sequenz B (CpG Island) aus der EMBL Datenbank

<http://srs.embl-heidelberg.de:8000/srs5/>

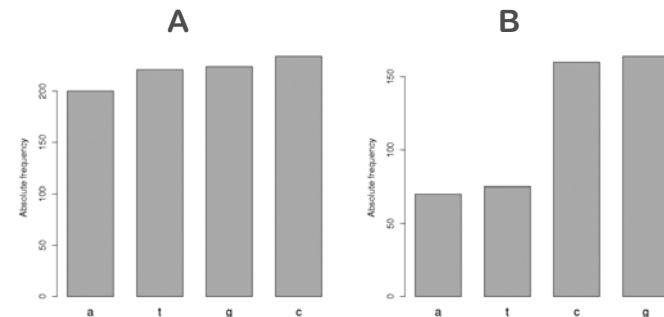
```
ID HSADRA2R standard; DNA; HUM; 3604 BP.
...
DE Human alpha 2 adrenergic receptor gene, complete cds.
...
FT source 1..3604
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /map="10q23-q25"
```

und der CpG Island Datenbank

<http://biomaster.uoi.no>

```
FT CpG island 1668..1996FT
/size=329FT
```

Zählen wir die a,c,g & t und machen einen Barplot



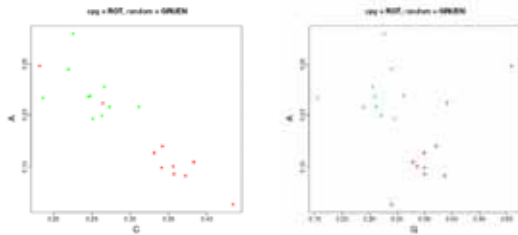
Wir haben absolute Anzahlen abgetragen

Es wird aber deutlich, daß in A alle vier Basen in etwa gleich häufig vorkommen, während in B die Basen c & g deutlich überrepräsentiert sind

A stammt aus einem Exon und **B** aus einem CpG Island

Samplen wir mehr Sequenzen aus Exons und CpG Islands ... jeweils 10

Und berechnen die relativen Häufigkeiten aller vier Basen



Im C-A Plot lassen sich die Gruppen nicht klar trennen

Im A-G Plot gibt es eine perfekte Trenngerade

Zusammenfassung der Beobachtungen

- CpG Islands enthalten ein höheren relativen Anteil an G & C's als Exons
- Oder, was das gleiche bedeutet: einen relativ geringeren Anteil an A & T's
- Durch Zählen des A und G Anteils lassen sich beide Sequenztypen leicht von einander unterscheiden

Zählen wir jetzt noch Buchstabenpaare wie AT, TT, GC ...

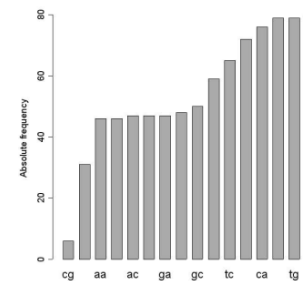
Wir zählen mit Überlappung:

z.B. In der Sequenz atatttt

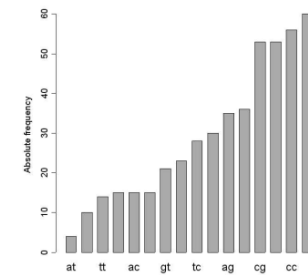
gibt es zweimal das Paar AT und dreimal das Paar TT

Für die beiden Sequenzen **A** & **B** ergibt sich

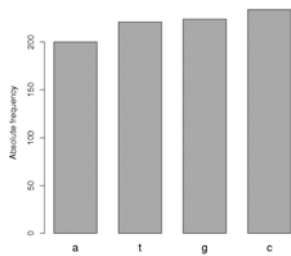
A



B



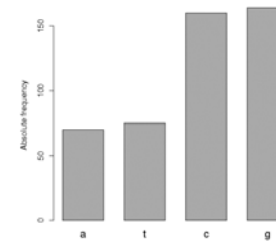
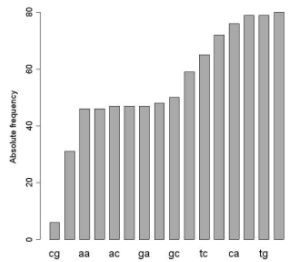
Das sieht sehr unterschiedlich aus



In Sequenz **A** fällt auf, daß obwohl die einzelnen Buchstaben (Basen) in etwa gleich oft vorkommen, dies nicht für die Buchstabenpaare gilt.

Auffällig ist, daß CG gemieden wird obwohl die einzeln Buchstaben C und G sogar am häufigsten sind.

Dies ist eine charakteristische Struktur der Sequenz. Ebenso wie auch das Nichtvorkommen der von Buchstabenfolgen wie `k1zrg` ein Zeichen von Struktur in natürlichen Sprachen ist.

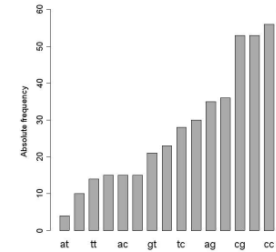


Auch in Sequenz **B** ist jede Menge Struktur zu erkennen, aber andere.

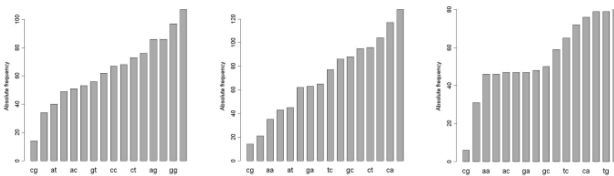
Hier spielt sich schon etwas auf der Ebene einzelner Buchstaben ab: C & G sind häufiger als A&T

Das Paar AT ist am Seltensten. Dies muß aber nicht auf eine Vermeidung dieses Paares zurückzuführen sein, sondern liegt daran, daß generell wenig A und T in der Sequenz sind. So sind Paare, die A und T enthaltenen generell seltener vertreten.

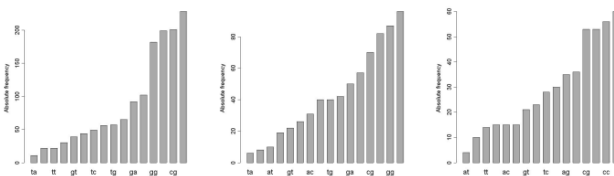
Im Vergleich zu Sequenz **A** fällt besonders auf, daß das Paar CG alles andere als selten ist.



Im Exon haben wir ein niedriges Auftreten von CG beobachtet, im CpG Island dagegen gerade nicht. Das hat System:



3 Exons



3 CpG Islands

Was sind Exons?

Sequenzabschnitte die transkribiert und nicht gespliced werden, die meisten davon kodieren für Proteine.

In Exons haben wir beobachtet, daß A,T,C & G in etwa gleich häufig sind, daß aber das Paar CG gemieden wird

Dies ist nicht nur typisch für Exons sondern für fast die ganze menschliche DNA

Unser Genom meidet die Kombination CG!

Unser Genom meidet die Kombination CG ...

... außer in relativ seltenen Abschnitten des Genoms:
den CpG Islands.

CpG Islands sind nicht transskribierte Sequenzabschnitte meist innerhalb von Promotorregionen, die gerade durch ihren hohen Anteil von C & G und insbesondere das relativ häufige Auftreten des Paares CG charakterisiert sind.

CpG Islands sind 300-3000 bp lang und nehmen etwa 1% des menschlichen Genoms ein.

Sie fallen gerade dadurch auf, daß sie sich von ihrer Buchstaben- und Buchstabenpaar-Zusammensetzung vom Rest des Genoms unterscheiden.

Ihre Entdeckung war ein statistischer Befund.

Wir haben zwei erstaunliche statistische Beobachtungen in genomischen Daten gemacht:

1. Das Genom meidet CG
2. Einzelne seltene Abschnitte tun dies gerade nicht

Was steckt dahinter?

Dem ist man nachgegangen ...

... und ist auf eine Zusammenhang mit Methylierungsmustern gestoßen.

DNA-Methylierung

- DNA-Methylierung ist eine Modifikation der DNA



- Beim Menschen werden nur Cytosine methyliert, und auch nur dann, wenn ein Guanin folgt ... also an CG-Positionen.

- Organismen, bei denen CG-Methylierungen auftreten, haben auch eine niedrige CG-Häufigkeit in ihren Genomen, und weisen CpG Islands auf.

- Beim Menschen sind ca. 70-80% der CG-Positionen methyliert ...

... aber gerade in den CpG Islands ist dies kaum der Fall

- Es ist beobachtet worden, daß die Methylierung in CpG Islands mit dem Alter der Menschen zunimmt.

- Methylierungsmuster (welche Stellen sind methyliert und welche nicht) werden nicht vererbt, entstehen aber früh im Embryo und sind danach relativ konstant.

- Liegt ein CpG Island im Promotorbereich eines Gens, so kann dieses Gen nur exprimiert (benutzt) werden, wenn keine Methylierung vorliegt

- Selten tritt Methylierung des Promotors ein und führt zum *Silencing* des Gens

- Dieses Silencing wird bei der Zellteilung vererbt und ist im wesentlichen irreversible (*epigenetic inheritance*)

- Es gibt Theorien, die dieses Phänomen mit dem Altern in Verbindung bringen

Promotor-Methylierung tritt nur bei vereinzelt Genen auf. Viele dieser Gene und ihr „Silencing“ sind mit der Entstehung von Krebs in Verbindung gebracht worden

Silencing führt zum Ausfall von Tumorsupressor-Genen

Gene	Map	Methylated
14-3-3 Sigma	1p	Breast and gastric cancers
ABL1 (P1)	9q34.1	50-100% CML, Some ALL
APC	5q21	Colon, gastric and esophageal
AR (Androgen Receptor)	Xq11-12	Prostate Cancer Cell Lines, Colon ACFs
BLT1 (Leukotriene B4 Receptor)		Various cell lines
BRCA1	17q21	10-20% Breast cancer, some ovarian
CALCA (Calcitonin)	11p15	25-75% Colon, lung, hematopoic neoplasms.
CASP8 (CASPASE 8)	2q33-34	Neuroblastoma
Caveolin	17q31.1	Breast cancer cell lines
CD44	11pter-p13	Prostate cancer
CFTR	7q31.2	Cell Lines

....

And this is only up to letter C !

<http://www3.mdanderson.org/leukemia/methylation/cgi.html>

Zusammenfassung der statistischen Begriffe:

- Häufigkeit
- Relative Häufigkeit
- Barplot
- Scatterplot
- Korrelation
- Scheinkorrelation
- Trenngerade