

Løsninger: Hvordan finder man en ny biomarkør?

I dette opgaveark skal du prøve at finde nye biomarkør-kandidater for Alzheimers sygdom. Du kan finde hjælp til opgaverne i afsnittet "Hvordan finder man en ny biomarkør?". Alzheimers sygdom er en neurodegenerativ sygdom, hvor der ophober sig proteiner i hjernen. Proteinerne klumper sig sammen og skader neuronerne, hvilket nedbryder hjernen gradvist.

1. Hvilke tre biomarkører kan lægerne måle i rygmarvsvæsken hos Alzheimers patienter, og hvilken slags biomarkører er de?

Beta-amyloid, tau og fosforyleret tau. De er alle tre proteiner og derfor molekylære biomarkører.

Det er et invasivt indgreb at tage en prøve fra rygmarvsvæsken. Det ville være hensigtsmæssigt, hvis lægerne kunne nøjes med at tage en blodprøve. Hvis proteiner og andre molekyler skal fra hjernen og ud i blodet, skal de passere blod-hjerne-barrieren. Blod-hjerne-barrieren beskytter hjernen mod uønskede stoffer udefra, men forhindrer samtidig passage fra hjernen og ud i blodet. Proteiner som beta-amyloid og tau kan ikke bare passere hen over blod-hjerne-barrieren. Det er altså ikke umiddelbart muligt at bruge dem som blod-baserede biomarkører.

I stedet for at måle på hele proteiner, har forskere prøvet at undersøge mængden af beta-amyloid peptider og tau peptider blodet.

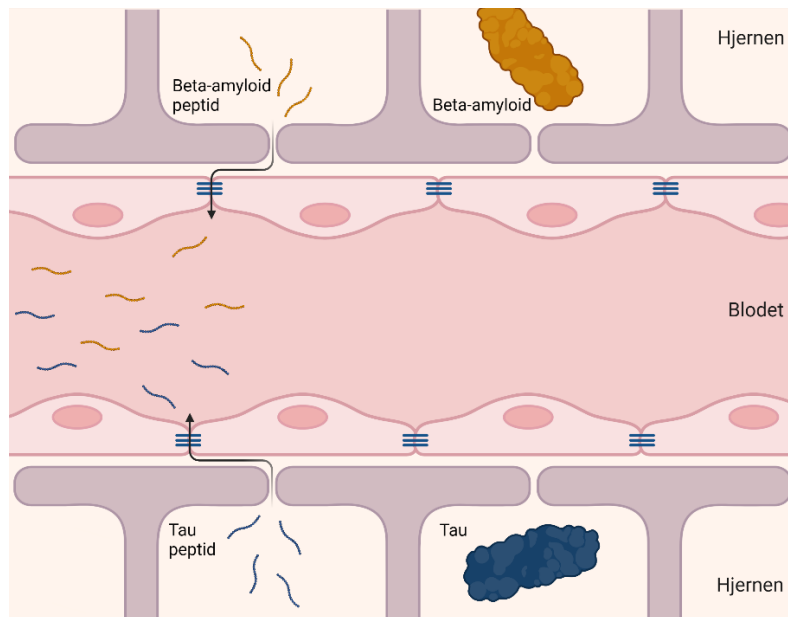
2. Hvad er et peptid?

Et peptid er en kort kæde af aminosyrer. Forskellen mellem proteiner og peptider er, at peptider er kortere end proteiner.

I hjernen findes der nogle enzymer, som kan klippe beta-amyloid og tau i mindre stykker til peptider.

3. Hvorfor tror du, at peptider potentielt er bedre biomarkør-kandidater end proteiner i forbindelse med hjernesygdomme? (Hint: Se figur 1)

Peptider er betydelig mindre end proteiner. Proteiner kan ikke passere blod-hjerne-barrieren, fordi de er for store. Peptider er meget mindre, og der er derfor større sandsynlighed for, at de kan passere fra hjernen og ud i blodet.



Figur 1: Blod-hjerne-barrieren.

I resten af opgaverne skal du prøve at finde frem til nogle biomarkør-kandidater for Alzheimers sygdom. Du kan downloade et Excel-ark her ([Link til hjemmesiden](#)), som indeholder data fra blodprøver taget fra 10 patienter med Alzheimers sygdom og 10 raske kontrolpersoner. Dataen er kun et lille udsnit af alle de peptidsekvenser, som forskerne har undersøgt.

4. Hvilken af de to primære måder at finde nye biomarkører på er blevet brugt til biomarkørforsøget? Forklar lidt om metoden.

Et opdagelses-baseret forsøg er blevet brugt til at finde nye biomarkør-kandidater. I opdagelses-baserede forsøg undersøger forskerne en helt masse peptider i stedet for at udvælge nogle fra starten. Forskerne har brugt proteomics til at finde ud af, om der er nogle af de mange peptider, som kunne være potentielle biomarkør-kandidater. De 18 peptider, som kan ses i Excel-arket, er kun en lille brøkdel af de peptider, som er blevet undersøgt i forbindelse med Alzheimers sygdom.

Start med at åbne Excel-arket og orientér dig om, hvad det indeholder. Nogle af kolonnerne er fyldt ud, mens andre kolonner er tomme. Du kommer til at fylde de tomme kolonner ud undervejs i opgaverne. Der er også et tomt plot (hvid firkant), som bliver fyldt ud, efterhånden som du får udfyldt kolonnerne. Det første, du kan kigge på, er kolonne B. Den indeholder 18 forskellige peptidsekvenser. Hvert bogstav svarer til en aminosyre i sekvensen.

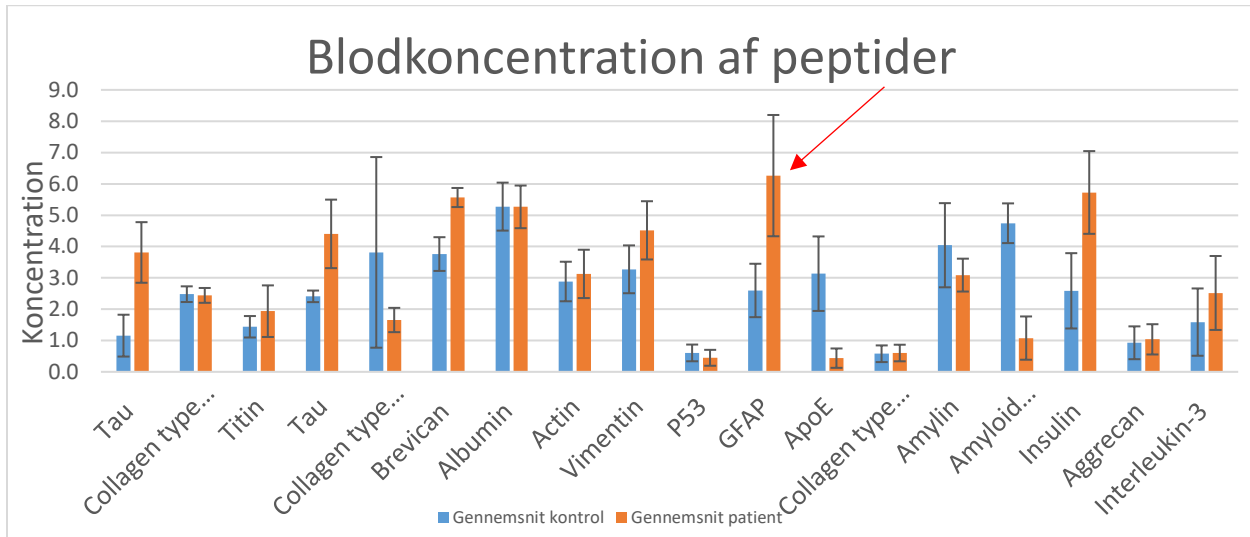
5. Hvilke aminosyrer består den første peptidsekvens, TPRGAAPPQ, af?

T	P	R	G	A	A	P	P	G	Q
Threonin	Prolin	Arginin	Glycin	Alanin	Alanin	Prolin	Prolin	Glycin	Glutamin

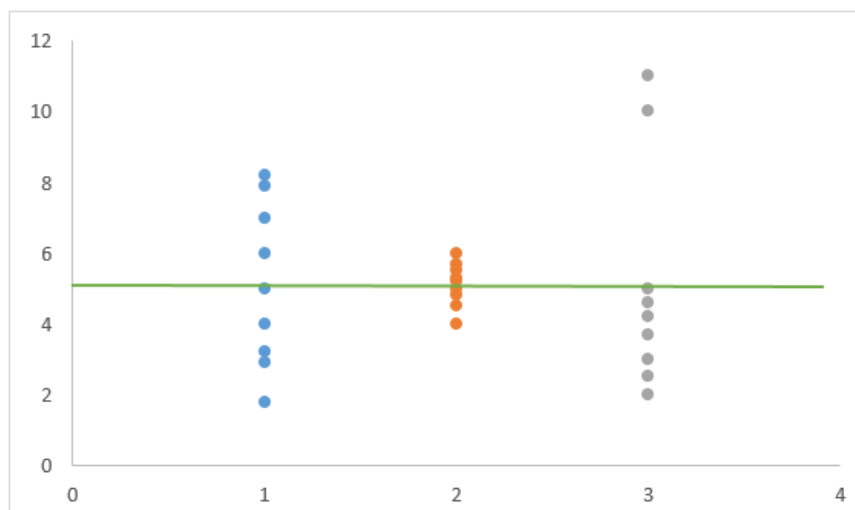
Det første, du skal gøre, er at finde gennemsnittet for hvert peptid i kontrolgruppen og patienterne. I kolonne X og kolonne Z skal du regne gennemsnittet ud for kontrolgruppen og patienterne. (Enheden på koncentrationen er vilkårlig, så den kan du se bort fra.)

6. Hvilket peptid har det højeste gennemsnit, og er det hos kontrolgruppen eller patienterne?
Hvad er gennemsnittet i den anden gruppe?

Peptid 11 har det højeste gennemsnit på 6,263 hos patienterne. Gennemsnittet for peptid 11 hos kontrolgruppen er 2,598.



Det er ikke nok kun at kigge på gennemsnittet. Det næste skridt er at regne standardafvigelse ud. Gennemsnittet fortæller ikke noget om, hvor spredte værdierne inden for hver gruppe er. Det er vigtigt at vide, når man skal finde ud af, om forskellen på grupperne er statistisk signifikant. Det vil sige, om man kan regne med, at forskellen i gennemsnittet ikke bare er tilfældig. Standardafvigelsen giver et indtryk af, hvor tætte alle værdierne i en gruppe er på hinanden. På figur 2 kan du se et eksempel på tre grupper, der alle har det samme gennemsnit, men har hver sin standardafvigelse.

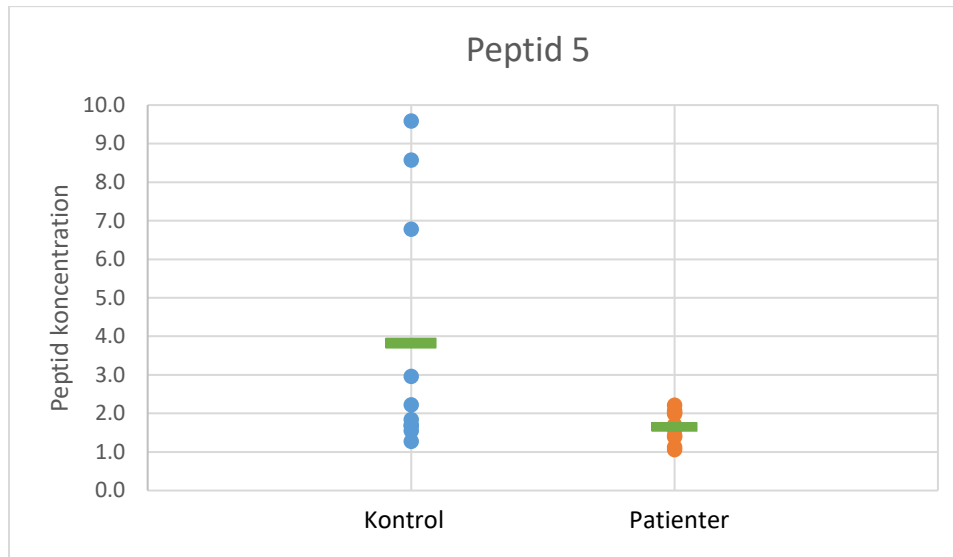


Figur 2: Gennemsnit og standardafvigelser. Gennemsnittet er markeret med en grøn streg og er ens for alle tre grupper.

Regn standardafvigelsen ud for kontrolgruppen i kolonne Y og patienterne i kolonne AA. (Hint: Hvis dit Excel-program er på dansk, kan du bruge funktionen =STDAFV())

7. Hvilket peptid har den største standardafvigelse? Giver det mening, når du kigger på de enkelte værdier i gruppen? Hvorfor?

Peptid 5 har den største standardafvigelse på 3,043 hos kontrolgruppen. Hvis man kigger på de enkelte værdier i kontrolgruppen for peptid 5, er de også meget spredte. Der er tre værdier i kontrolgruppen, som er meget højere end de andre. De medfører en stor standardafvigelse for kontrolgruppen. På nedenstående plot kan du se, den store variation i koncentrationen af peptid 5 for kontrolgruppen. Gennemsnittet for kontrolgruppen og patienterne er markeret med en grøn streg.



Nu har du regnet alle værdierne ud, som du skal bruge for at løse opgaven. Der skulle gerne være dukket et flot plot op undervejs. Plottet er et søjlediagram, der viser den gennemsnitlige blodkoncentration af de 18 peptider i kontrolgruppen og i patienterne. De sorte tynde streger i toppen af hver søjle er standardafvigelse.

8. Inspicér plottet. Hvilket peptid ligner det har den mindste koncentrationsforskel imellem kontrolgruppen og patienterne? Hvilket peptid har den største?

Ud fra plottet ser det ud til, at den mindste koncentrationsforskel er for peptid 2, 7 eller 13. Den største koncentrationsforskel er for peptid 11 eller peptid 15.

Udvælg tre til fem peptider, som du tror, kunne være gode biomarkør-kandidater.

9. Hvilke peptider har du valgt, og hvorfor har du valgt dem?

Der er flere peptider, som kan vælges. Et fornuftigt kriterie er, at forskellen i gennemsnittet mellem kontrolgruppen og patienterne er relativt stort. Derudover skal standardafvigelse helst ikke overlappes på plottet. Hvis de overlapper, er der for stor variation i hver gruppe, og det bliver usikkert, om gennemsnittene reelt er forskellige, eller om det bare er tilfældigt. De to kriterier opfylder de syv peptider i nedenstående tabel. I virkeligheden ville forskerne bruge statistiske test til at udvælge de bedste biomarkør-kandidater, men vi holder os til disse to kriterier.

Peptid	Peptidsekvens	Højeste koncentration
1	TPRGAAPPGQ	Patienterne
4	SSTGSIDMVD	Patienterne
6	DVLEGDSSSED	Patienterne
11	ESLEEEIRFL	Patienterne
12	ELRQQTEWQS	Kontrolgruppen
15	WHSFGADSVF	Kontrolgruppen
16	KFSYIRTSFD	Patienterne

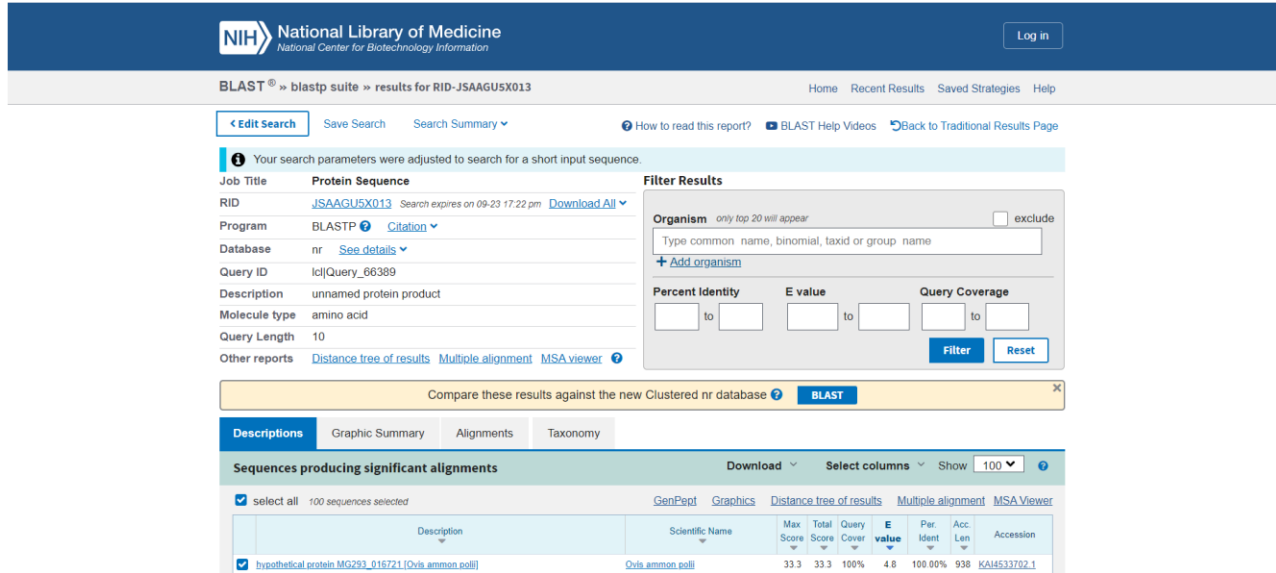
10. Er koncentrationen af dine peptider højere eller lavere i Alzheimers patienterne sammenlignet med kontrolgruppen?

Se tabellen ovenover fra spørgsmål 9.

Ind til nu har du kun arbejdet med peptidsekvenser, men det kunne jo være interessant at finde ud af, hvilke proteiner peptiderne kommer fra. Det kan man gøre med programmet Protein BLAST (BLASTp). BLASTp kan sammenligne en hvilken som helst peptidsekvens med sekvenser fra en proteindatabase. Den peptidsekvens, som du gerne vil sammenligne med databasen, kaldes en query-sekvens. Når du har indsat din query-sekvens i programmet, søger den efter lignende sekvenser i databasen. Det kaldes at "blaste" sekvensen.

Gå ind på BLASTp via linket <https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>. Indsæt den ene af de peptidsekvenser, som du har udvalgt. Den eneste parameter, du skal ændre på, er feltet "Organism". Alle peptiderne stammer fra blodprøver fra mennesker, så specificér hvilken organisme din query-sekvens kommer fra ved at skrive "Homo sapiens (taxid:9606)". Tryk til sidst på BLAST nede i bunden til venstre for at "blaste" din query-sekvens. Det kan godt tage et par minutter, før du får et svar. Du kan lære meget mere om BLAST [her](#), hvis du er interesseret.

Når du får et svar, ser hjemmesiden nogenlunde ud som på figur 3. Der står rigtig mange forskellige oplysninger, men det er kun nogle af dem, som er interessante for den her opgave. BLASTp viser de 100 bedste hits, som du kan se, når du scroller ned på hjemmesiden. Hvert hit er en proteinsekvens, der i større eller mindre grad matcher din query-sekvens.



Figur 3: BLASTp hit.

Der er tre tal, som kan være gode at kigge på: **Max-score, query-cover og E-value.**

- Max-score er den score, som hittet har opnået.
- Query-cover er den procentvise overensstemmelse mellem din query-sekvens og hittet. Hvis den er 100%, er alle aminosyrerne ens.
- E-value fortæller noget om, hvor unikt hittet er. E-value angiver, hvor stor sandsynligheden er for at få et hit med samme score på tilfældigvis. Jo mindre E-value desto bedre. Så er hittet meget unikt, og der er større chance for, at matchet er rigtigt.

Blast alle de peptidsekvenser, som du har valgt, og udfyld tabellen. Udfyld også kolonne C i Excel-arket med proteinnavnet på de peptider, som du har undersøgt.

Peptid	Peptidsekvens	Protein ("Description")	Max-score	Query-cover	E-value
1	TPRGAAPPGQ	Tau	33,3	100%	0,005
4	SSTGSIDMVD	Tau	34,1	100%	0,002
6	DVLEGDSSSED	Brevican	33,7	100%	0,003
11	ESLEEEIRFL	Glial fibrillary acidic protein (GFAP)	35,8	100%	0,0006
12	ELRQQTEWQS	Apolipoprotein E	37,5	100%	0,0001
15	WHSFGADSVF	Beta-amyloid	36,3	100%	0,0004
16	KFSYIRTSFD	Insulin receptor	36,3	100%	0,0004

11. Hvilke proteiner tilhører dine peptidsekvenser?

Tau, Brevican, GFAP, Apolipoprotein E, beta-amyloid, Insulin receptor

12. Er proteinerne relevante i forhold til Alzheimers sygdom? Hvorfor/hvorfor ikke? Hvis du har fundet et protein, som du ikke kender, så prøv at slå det op på nettet for at finde ud af noget om dets funktion og lokalisation i kroppen. Du kan for eksempel kopiere navnet og slå det op på UniProt via linket <https://www.uniprot.org/>.

Tau er et af de proteiner, som man ved er involveret i Alzheimers sygdom. Tau findes inde i neuronerne, hvor de stabiliserer neuronernes cytoskelet. I Alzheimers patienter bliver tau fosforyleret, hvilket destabiliserer neuronernes cytoskelet og ødelægger dem.

Brevican findes på overfladen af neuronerne i hjernen. Man mangler stadig viden om brevican, men forskning tyder på, at det er involveret i beskyttelsen af neuroner, og neuronernes evne til at danne nye forbindelser i hjernen. Brevican kunne derfor godt være relevant i forhold til Alzheimers sygdom, hvor neuronerne bliver nedbrudt.

GFAP findes bl.a. i astrocytter i hjernen, hvor det er en del af disse cellers cytoskelet. Når hjernen bliver skadet, producerer astrocytterne mere GFAP. GFAP kunne derfor godt være relevant i forhold til Alzheimers sygdom, hvor neuronerne bliver ødelagt.

Apolipoprotein E er et protein, som binder fedtstoffer (lipider) og transporterer dem rundt i kroppen. Proteinets er også involveret i at fjerne beta-amyloid fra hjernen. Et dansk studie har vist, at personer med et lavt niveau af Apolipoprotein E i blodet har større risiko for at udvikle demens og Alzheimers sygdom. Det kan du læse mere om [her](#), hvis du er interesseret.

Beta-amyloid er et af de proteiner, som man ved er involveret i Alzheimers sygdom. Beta-amyloid klumper sig sammen og danner plaques i hjernen, hvilket forhindrer kommunikationen imellem neuronerne.

Insulin receptorer er vigtige for, hvordan kroppens celler optager sukker. Hos diabetes patienter er der problemer med cellernes sukeroptag. Det kan både skyldes selve insulin produktionen og insulin receptorenes respons på insulin. Diabetes øger risikoen for at udvikle Alzheimers sygdom. Det er muligvis derfor, at der er forskel mellem insulin receptor koncentrationen hos patienterne og kontrolgruppen. Insulin receptorer er dog ikke specifikke for hjernen eller Alzheimers sygdom, og det bliver derfor svært at bruge dem som biomarkører.

13. Hvad ville næste skridt være, hvis du skulle undersøge dine biomarkør-kandidater nærmere?
(Hint: Tænk på de fire trin i biomarkør-opdagelsen)

Efter forskerne har fundet nogle biomarkør-kandidater, skal de validere deres resultater. Det næste skridt er altså at undersøge de udvalgte biomarkør-kandidater i et større studie med flere prøver, for at se om de får samme resultat.

Forestil dig, at dine biomarkør-kandidater er så gode til at diagnosticere Alzheimers patienter, at de er blevet godkendt til brug i sundhedsvæsenet. De er dog stadig ikke lige så præcise som biomarkørerne fra rygmarvsvæsken. Dine biomarkører bliver derfor brugt som en form for "førstesortering", hvor patienterne med positiv biomarkør-status bliver sendt videre i systemet for at få taget en prøve fra rygmarvsvæsken.

14. Hvis blod-baserede biomarkører skal fungere som en "førstesortering", hvad er så vigtigst: At de har en høj sensitivitet eller en høj specificitet? Hvorfor?

Det er vigtigst, at biomarkørerne har en høj sensitivitet, så alle de syge bliver sendt videre i systemet. Det er okay, at nogle af de raske også sendes videre, så længe de bliver sorteret fra i de videre undersøgelser. I en førstesortering er en høj sensitivitet altså vigtigere end en høj specificitet.