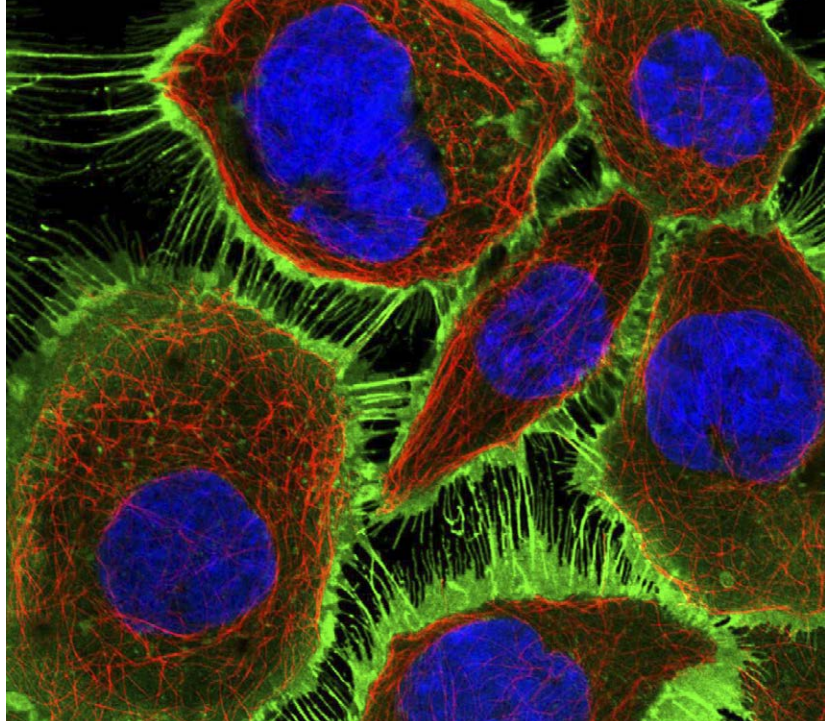


EGFR (epidermal growth factor receptor), det gröna på bilden, visas med hjälp av konfokalmikroskop och immunofluorescence i en cellinje med ursprung från hudcancer. Receptorn är lokaliserad till plasmamembranet (de långa trådarna är så kallade membranutskott) och reglerar celledelning. Mikrotubuli, en del av cellskelettet, visas i rött och cellkärnan i blått.

Källa: Emma Lundberg, Cell Profiling, Science for Life Laboratory
Bilden finns på The Human Protein Atlas, www.proteinatlas.org.



The Human Protein Atlas

För ungefär ett år sedan lanserades resultatet av ett av de största svenska vetenskapliga projekten genom tiderna – en proteinatlas där en stor del av människans proteiner, som kodas av omkring 20 000 gener, kartlagts. Proteinatlasen är till nytta för såväl forskare som läkemedelsföretag men den kan även komma till användning på flera sätt i skolan.

Text: Per-Henrik Edqvist, Tissue Profiling Facility, Science for Life Laboratory
E-post: per-henrik.edqvist@igp.uu.se

Kring millennieskiftet föddes en idé hos forskare på Kungliga Tekniska Högskolan och Uppsala universitet. Den mänskliga arvmassan hade nyligen kartlagts och publicerats i USA och nästa naturliga steg var att börja kartlägga var i kroppen arvmassans olika gener uttrycks. Det vill säga, var finns de proteiner som generna kodar för och hur skiljer sig olika vävnader och celltyper från varandra med avseende på dessa proteiner? Idén blev grunden för "The Human Protein Atlas Project" och möjliggjordes genom bidrag från Knut och Alice Wallenbergs Stiftelse. Mer än ett decennium senare, och med över 1 000 manårs arbete bakom sig, publicerades i januari 2015 det första utkastet av kartläggningen av människans proteiner i normala, friska vävnader i tidskriften Science. Den mänskliga proteinatlasen finns tillgänglig gratis på nätet för vem som helst att använda på www.proteinatlas.org. Förutom att rikta sig till vetenskapssamhället är atlasen även en utmärkt resurs i undervisning om kroppen, cellbiologi och molekylärbiologi.

Vad är proteiner?

Proteiner är "maskiner" som har olika funktioner i våra celler och vävnader. De tillverkas i och av alla kroppens celler och instruktionerna för hur varje enskilt protein ska byggas, i vilken ordning

aminosyror ska sättas ihop, finns lagrat i våra gener. Instruktionen går dock inte direkt från DNA till protein utan mRNA-molekyler fungerar som en mellanhand genom att kopiera DNA:ts instruktioner och sedan ta sig till cellens "proteinfabriker", ribosomerna. På ribosomerna sätts aminosyror samman enligt mRNA-instruktionen och bildar därmed ett specifikt protein.

Människan har cirka 20 000 olika proteinkodande gener. Många gener kan dock ge upphov till flera varianter av samma protein så teoretiskt sett kan *minst* 20 000 olika möjliga proteiner/proteinvarianter bildas. Alla 20 000 proteinkodande gener är dock inte uttryckta samtidigt utan olika typer av celler uttrycker olika proteiner. Vissa proteiner är nödvändiga för alla cellers normala basalfunktioner och finns därmed i princip alltid uttryckta i alla kroppens celler. De kan till exempel sköta reparationer, energiomsättning eller proteintillverkning. Andra proteiner sköter specialiserade funktioner som är karaktäristiska för en viss celltyp. Leverceller uttrycker exempelvis proteiner som kan ta hand om och bryta ner gifter, medan hudceller uttrycker proteiner som skapar en barriär mot omvärlden. Nervceller är specialiserade på att ta emot och skicka elektriska impulser och tarmceller på att ta upp näring och vätska från födan.

Se en pdf-version av artikeln med klickbara länkar på vår hemsida.

Hur går kartläggningen till?

När man talar om genuttryck kan man mena både förekomsten av ett specifikt mRNA (varje mRNA kommer från en specifik gen) eller avse förekomsten av ett specifikt protein (som kommer från ett specifikt mRNA). Ibland behöver man dock specificera om man menar förekomsten av mRNA (mRNA-uttryck) eller förekomsten av protein (proteinuttryck). Proteinatlasen använder sig av två olika, men komplementära, metoder för att mäta och påvisa genuttryck. Dels *immunhistokemi*, då antikroppar binder till olika proteiner, dels *RNA-seq*, då mRNA-nivåer mäts med hjälp av RNA-sekvensering.

Det som gör proteinatlasprojektet unikt i sitt slag är den storskaliga och strömlinjeformade produktionen av antikroppar och användandet av vävnadsprover för att påvisa i vilka vävnader som olika proteiner finns. Antikroppar har den generella förmågan att binda till ett visst protein med hög specificitet. Genom att skapa tiotusentals olika antikroppar, som var och en kan binda till ett specifikt protein, har projektet kunnat kartlägga uttrycket av nästan alla människans proteiner i ett stort antal vävnadstyper med hjälp av immunhistokemi (figur 1). Antikroppar tillsätts till tunna vävnadssnitt, varpå de binder in till sitt målprotein och fastnar. Kemiska reaktioner skapar därefter en synlig färg där antikroppen bundit till proteinet och i ett mikroskop kan man sedan analysera i vilka vävnader och celltyper som proteinet finns eller inte finns. I proteinatlasen finns över 13 miljoner högupplösta digitala bilder av immunhistokemiskt färgade prover tillgängliga för vem som helst att själv studera vid sin dator, som om man tittade i ett mikroskop.

Fördelen med immunhistokemi är att man får en direkt, intuitiv och synlig av-på-reaktion, som visar var proteinet finns i en komplex väv-

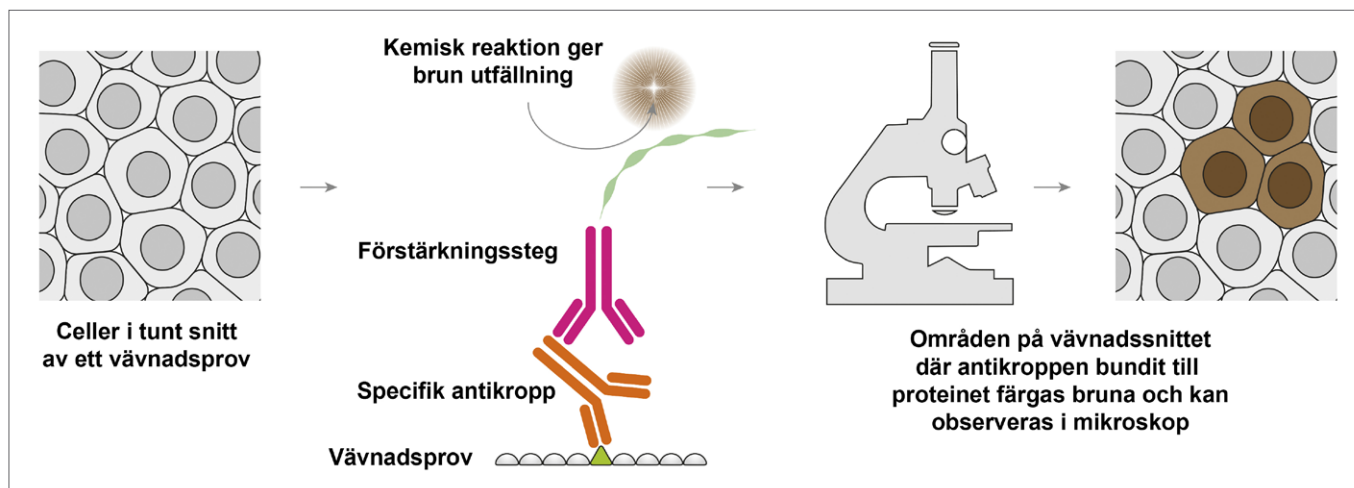
nad som oftast består av en blandning av flera celltyper. Man kan se om olika celltyper i samma vävnad uttrycker proteinet olika mycket och får även en grov uppskattning av den subcellulära förekomsten av proteinet, det vill säga om proteinet finns i kärnan, membranet och/eller cytoplasman. Nackdelen med immunhistokemi är att det oftast krävs ett mått av vävnadskunnande för att tolka det man ser på ett korrekt vis och kunna separera olika celltyper från varandra. En annan nackdel är att det är en kvalitativ metod som inte berättar hur mycket av proteinet det finns i olika celler, utan endast kan ge en grov kvantitativ information baserad på relativa skillnader i färgens styrka mellan olika celler.

Fördelen med att analysera mRNA genom sekvensering (RNA-seq) är att man får en exaktare, kvantitativ, mätning av antalet mRNA-molekyler som finns i provet, vilket man sedan kan jämföra med andra prover. Man får samtidigt i en enda analys veta vilka av nästan alla 20000 proteinkodade gener som uttrycks i ett och samma prov. Med immunhistokemi måste man jobba med ett protein i taget. Nackdelen med RNA-seq är att man inte får veta vilka celler mRNA-molekylerna kommer ifrån, eller i vilken utsträckning vissa celltyper bidrar mer eller mindre. Man får istället ett genomsnitt av hela provet utan cellulär upplösning.

Dessa två metoder skiljer sig alltså åt på ett antal viktiga punkter och ger olika sorters information. Men summerar man bägge får man en mycket bra bild av vilka gener som uttrycks i våra kroppar, vävnader och olika sorters celler.

Att använda proteinatlasen

Proteinatlasen består av fyra olika delatlasar: en *vävnadsatlas* som visar proteinuttrycket i 44 av kroppens normala vävnader med en upplösning



Figur 1: Principen för immunhistokemi. Antikroppar (orange) riktade mot ett specifikt protein tillsätts till ett tunt vävnadssnitt, och genom kemiska reaktioner kan sedan proteinets förekomst påvisas genom en färgreaktion som är synlig i mikroskop.

Bildkälla: The Human Protein Atlas, www.proteinatlas.org

på 83 olika celltyper, en *subcellulär atlas* som med hjälp av högupplöst konfokalmikroskopi visar i vilka av cellens organeller som proteinerna finns, en *cellinjeatlas* som visar vilka proteiner som finns i 46 på forskningslaboratorier vanligt förekommande cellinjer och en *canceratlas* som visar proteinuttrycket i de 20 vanligaste förekommande formerna av cancer. Canceratlasens syfte är bland annat att utgöra en startpunkt för forskare som letar efter nya intressanta så kallade biomarkörer, som kan användas inom sjukvård och forskning. Ett flertal sådana markörer har redan identifierats tack vare proteinatlasprojektet.

Som tidigare nämnts finns proteindata med bilder för de flesta normalvävnader och cellinjer integrerat med mRNA-data i databasen. Allt är sökbart och på förstasidan kan man snabbt och enkelt söka efter enskilda proteiner, till exempel insulin, som uttrycks i de Langerhanska öarna i bukspottkörteln. Genom en avancerad sökning kan också exempelvis en diabetesforskare få en lista över gener som är starkt uttryckta i de Langerhanska öarna, men inte i resten av bukspottkörteln eller andra närliggande organ. Vid sidan av rena sökfunktioner finns även en rad mycket användbara och användarvänliga sidor som möjliggör interaktiv utforskning av kroppens olika organ och deras likheter och skillnader.

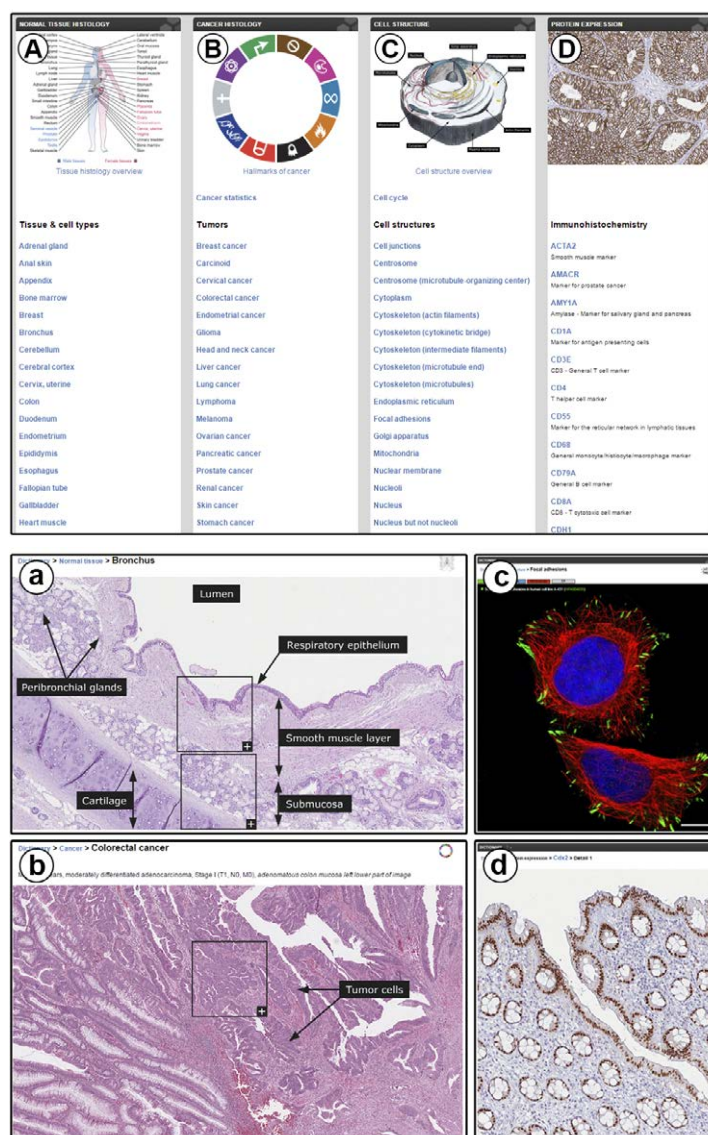
Proteom och Dictionary

Ett *proteom* är summan av de proteiner som uttrycks i ett definierat område vid en specifik tidpunkt, till exempel i en cell, i ett organ eller i en hel individ. Genom att kategorisera proteinuttrycket i olika vävnader kan man alltså definiera olika proteom, till exempel hudproteomet, leverproteomet, och så vidare. På proteinatlasen kan man enkelt utforska och jämföra olika organs proteom i sektionen *det vävnadsspecifika proteomet*. Man kan där även hitta och utforska de "specialistgener" som är gemensamma för till exempel huden, matstrupen och tonsillen. Därtill finns texter och exempelbilder för varje organ och dess särdrag som tydliggör kopplingen mellan proteinuttryck och specifika funktioner i en viss vävnad.

För den molekylärbiologiskt intresserade finns sektioner där man kan läsa om olika molekylärbiologiska metoder där antikroppar används. Det är även möjligt att interaktivt utforska olika delar av vårt proteom, till exempel transkriptionsfaktorer eller sekretoriska proteiner. Man kan även undersöka vilka gener som vanligtvis är relaterade till olika former av cancer i *cancerproteomet* eller studera *läkemedelsproteomet*, som listar de proteiner som är måltavlor för läkemedel.

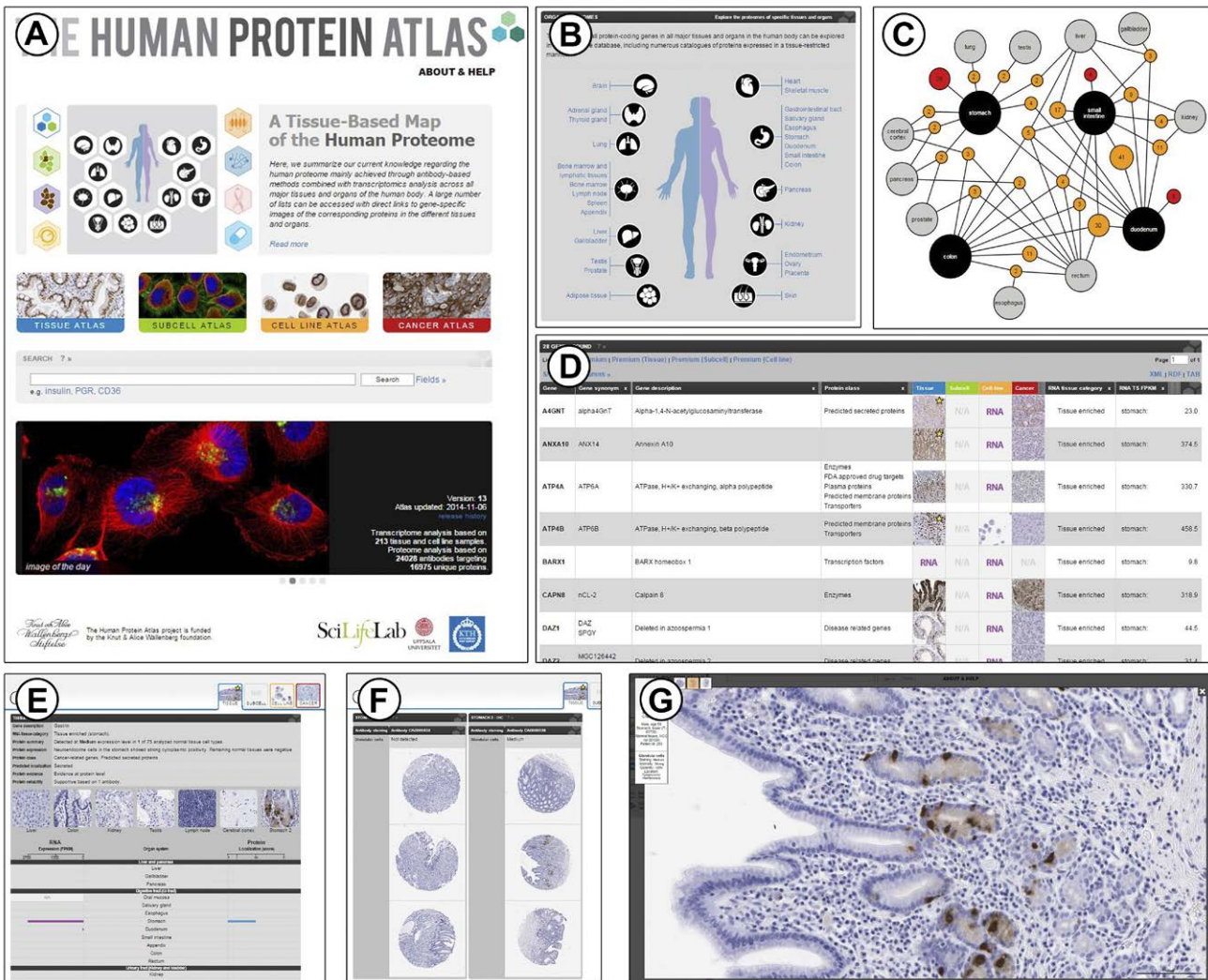
Proteinatlasen har också en sektion som heter *Dictionary* (figur 2) som är skapad för att ge be-

sökarna en introduktion i vävnadskunskap (histologi). Här beskrivs och visas olika vävnader och organ som de ser ut i mikroskop (A, a). Syftet är att ge besökare bakgrundkunskap för att själva kunna förstå och tolka de immunhistokemiskt färgade bilderna, men Dictionary kan även användas som en resurs i undervisningssyfte. Förutom vävnadsbaserad normal histologi finns där information hur olika tumörtyper ser ut, tillsammans med information och statistik som rör cancer (B, b). Det finns också en sektion som handlar om cellens organeller och deras funktioner, med bilder som visar hur olika organeller ser ut i mikroskop (C, c). Man kan också undersöka hur vanliga diagnostiska antikroppar som används inom sjukvården ser ut i olika vävnader (D, d).



Figur 2: I Dictionary-avsnittet kan man bland annat lära sig mer om vävnader och hur de ser ut i mikroskop. Man kan studera normala vävnader (A, a), olika former av cancer (B, b), cellens organeller (C, c) och hur vanliga diagnostiska antikroppar som används inom sjukvården ser ut i olika vävnader (D, d). Bild a–d visar exempel på vad som förekommer under de olika rubrikerna, till exempel normal luftstrupe (a), tjocktarmscancer (b), två celler med fokala celladhesioner (c, grön färg) samt färgning av det tarmspecifika proteinet CDX2 i normal tjocktarm (d, brun färg).

Bildkälla: The Human Protein Atlas, www.proteinatlas.org



Figur 3: Exempel på hur informationen i proteinatlasen (A) kan utnyttjas genom att börja på organnivå (B) och därefter stegvis (C–F) närm till högupplösta bilder på hur enskilda proteiner uttrycks i olika vävnader (G). Bilden i G visar uttrycket av proteinet gastrin (GAST) i normal magsäck (brunfärgade celler). De celler som inte uttrycker gastrin är blå. Gastrin är ett lokalverkande hormon som bara finns i en specialiserad typ av magsäcksceller (de så kallade G-cellerna) och när det frigörs stimulerar det frisättningen av saltsyra från en annan specialiserad celltyp i magsäcken, de så kallade parietalcellerna.
Bildkälla: The Human Protein Atlas, www.proteinatlas.org

Övningsexempel för skolan

Mag-tarmkanalen (på engelska: gastrointestinal tract) består av flera olika organ, bland annat magsäcken, tolvfingertarmen, tunntarmen och tjocktarmen. Eftersom dessa organ är närbesläktade kan man anta att de även är lika med avseende på de "specialistgener" de uttrycker. Följande övning erbjuder en möjlighet att undersöka både skillnader och likheter mellan dessa fyra organ.

Gå in på proteinatlasens hemsida (figur 3A), www.proteinatlas.org, och klicka på bildlänken *A Tissue-Based Map of the Human Proteome* och därefter på länken *Gastrointestinal tract* (figur 3B) för att komma till sidan som beskriver det mag-tarmspecifika proteomet. Börja övningen med att läsa på om funktionen hos de olika delarna som ingår i mag-tarmsystemet, inklusive spottkörtlarna och matstrupen. Använd informationen på sidan samt länkar till de olika organen som finns i en grå box till höger om texten.

1. Beskriv översiktligt vad de olika organen har för uppgifter (cirka 1–2 meningar per organ).

2. Vad är/gör ett enzym (använd internet eller böcker för att ta reda på svaret)?

3. Två proteiner som nämns är GAST och GIF. Vad är dessa proteiners roll i matsmältningen?

Långt ner på gastrointestinal-sidan finns ett "cytoscape-diagram" (figur 3C) som med olika cirkular visar hur många högt uttryckta specialistgener som finns i varje organtyp (röda) och hur många som delas mellan olika organ (gula). Cirkularna är klickbara och tar användaren till en lista över de gener som cirkeln representerar.

4. Vilket av följande organ har flest högt uttryckta specialistgener; magsäcken (stomach), tolvfingertarmen (duodenum), tunntarmen (small intestine) och tjocktarmen (colon)? Klicka på den röda cirkeln som visar antalet specialistgener i organet och gå till nästa punkt.

5. Du ser nu en lista över de gener som är högt uttryckta specifikt i detta organ (figur 3D). Hitta till exempel proteinerna som heter GAST och GIF i listan och klicka på något av dessa för att undersöka hur uttrycket ser ut i kroppens olika organ. Gå till nästa punkt.

6. Du ser nu en lista över vävnader med blå och lila liggande staplar bredvid varje organ (figur 3E). De blå staplarna visar proteinuttrycket som påvisats med immunhistokemi och de lila staplarna visar motsvarande mRNA-nivå för genen i respektive organ. Ibland är dessa staplar inte helt överensstämmande, vilket kan bero på att vissa antikroppar har en tendens att även känna igen andra proteiner än just det de ska vara specifika mot. Man kan klicka på vilket organ som helst i listan för att se de immunhistokemiskt färgade bilderna (figur 3F) och navigera i dem med musen som om man tittade i mikroskop (figur 3G). Titta och jämför hur proteinuttrycket ser ut på bilderna för GAST och GIF. Skiljer det sig åt? Finns proteinerna i olika celltyper?
7. Välj ett eget protein i listan från punkt 5 (till exempel LIPF) och ta själv reda på mer om detta proteins funktion i kroppen, genom att till exempel läsa på GENE/PROTEIN-fliken (figur 3E) eller söka på nätet.
8. Gå tillbaka till cytoscape-diagrammet (figur 3C). Hur många högt uttryckta gener delar endast tolvfingertarmen och tunntarmen? Klicka på den aktuella cirkeln och titta närmre på proteinet DEFA5. Detta protein uttrycks bara av en särskild celltyp som finns i dessa organ. Vad heter den celltypen och vad har den för funktion? Läs på GENE/PROTEIN-fliken och sök eventuellt mer information på nätet.

Svar på frågorna i övningen finns på Bioresurs hemsida, www.bioresurs.uu.se.

Övriga förslag på tillämpningar

- *Proteiner*: Använd läroboken för att hitta olika proteiner som nämns i samband med till exempel matsmältningen (exempelvis amylas och insulin), muskelsystemet (aktin och myosin) eller energiomsättningen (ATP-syntas). Sök upp dem i proteinatlasen. Hur och var uttrycks de? Hur är detta relaterat till funktion? Finns proteinerna överallt eller är de specifika för vissa organ? Tänk på att proteinnamnen behöver skrivas på engelska i sökningen samt att det ofta finns flera proteiner med likartade namn som kommer från olika gener.
- *Människokroppen*: Använd de olika organsidorna under *Organ proteomes* (figur 3B) för att få information om olika organs specialiserade funktioner och se exempel på genuttrycksnivå. Det finns mycket text och många exempel med bilder att använda för att hitta inspiration eller material till undervisning.
www.proteinatlas.org/humanproteome
På www.proteinatlas.org/download/poster_proteome.pdf finns en sammanfattande plansch med titeln *The human proteom*
- *Cellbiologi*: Studera hur cellens organeller ser ut och fungerar (figur 2C). Vad är en nukleol? Vad skiljer det endoplasmatiska retiklet från golgiapparaten? Vad skiljer kärnmembranet från plasmamembranet? Genom högupplösta konfokalmikroskopiska bilder kan man se hur dessa organeller ser ut och var de befinner sig inuti cellen.
www.proteinatlas.org/learn/dictionary/cell
- *Histologi*: Lär om kroppens funktioner genom att studera hur olika vävnader ser ut i mikroskop (figur 2A). Hur ser njurens glomeruli ut och vad gör de? Hur skiljer sig utseendet hos lillhjärnan (cerebellum) från storhjärnan (cortex) och vad har de olika celltyperna för funktioner?
www.proteinatlas.org/learn/dictionary/normal

Som komplement till punkterna om cellbiologi och histologi finns även en nedladdningsbar manual på 84 sidor som visar och beskriver olika immunhistokemiskt färgade vävnader och celler samt hur cellens olika organeller kan se ut i vanligt ljusmikroskop.
www.proteinatlas.org/download/annotation_dictionary_print.pdf (95 MB)
www.proteinatlas.org/download/annotation_dictionary_screen.pdf (15 MB)
- *Molekylärbiologi*: Antikroppar är centrala redskap i många molekylärbiologiska metoder inom den moderna forskningen. På proteinatlasen kan man hitta information om hur exempelvis immunhistokemi, flödescytometri, proximitetsligering eller Western blotting går till.
www.proteinatlas.org/learn/method
- *Diagnostik*: Antikroppar är oundgängliga redskap inom sjukvården när det gäller att ställa diagnoser baserade på vävnadsprover. Det finns ett hundratal väl etablerade standard-antikroppar som används för olika syften inom både sjukvård och forskning. Ta en titt på ett urval av dessa för att se hur de ser ut och hur de används kliniskt (figur 2D).
www.proteinatlas.org/learn/dictionary
- *Cancerbiologi*: Cancer är ett samlingsnamn på en mängd olika sjukdomstillstånd som har gemensamt att kroppens egna celler växer okontrollerat. På proteinatlasen kan man läsa mer om olika former av cancer och se hur de ser ut i mikroskop. Man kan även studera *The cancer proteome* för att få en introduktion till gener som har visat sig viktiga för uppkomsten av cancer i olika organ.
www.proteinatlas.org/learn/dictionary (figur 2B), www.proteinatlas.org/humanproteome/cancer