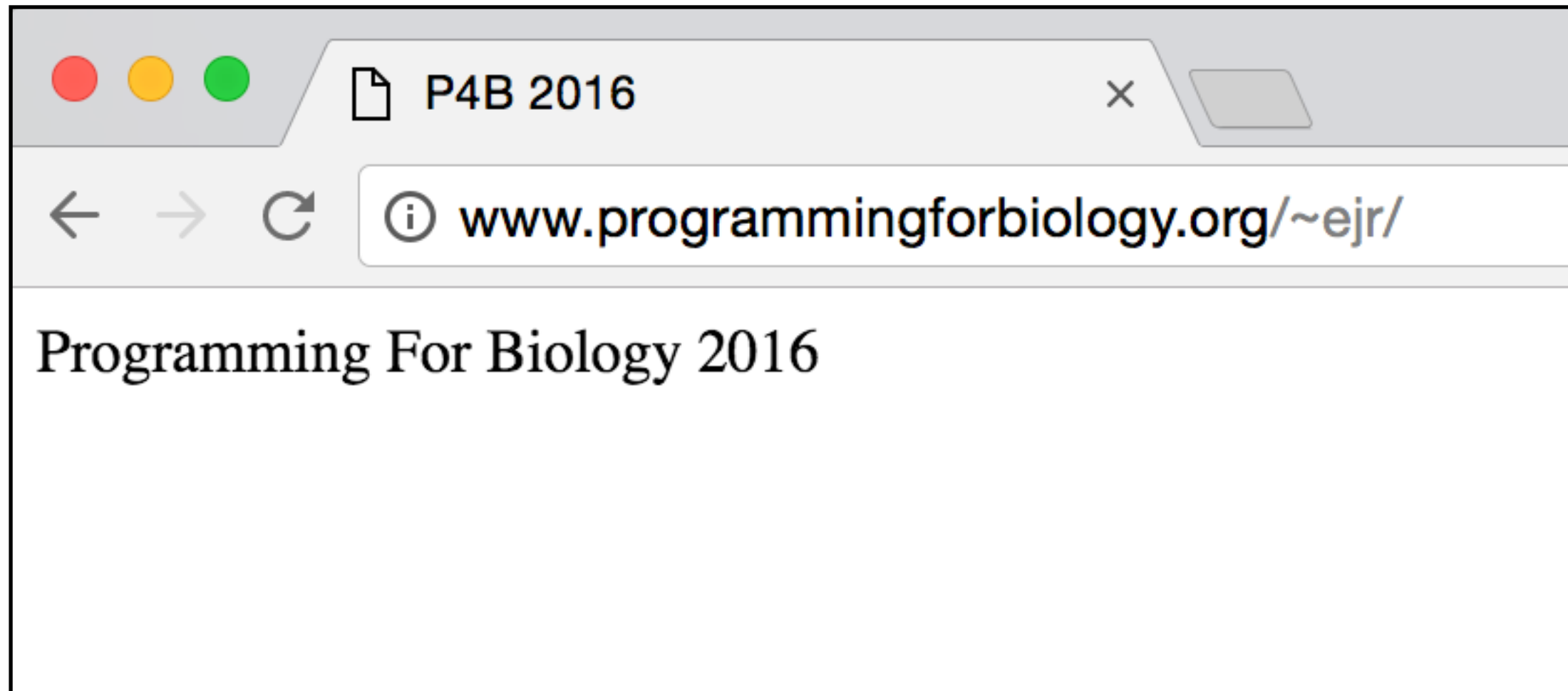`<html>`

<html>

- "HyperText Markup Language"
- Not a programming language
- Stored in text files (just like Perl)
  - On a server
  - Downloaded, interpreted and displayed by a web browser (e.g. Chrome or Firefox)

A basic page

A basic page

```html
<!doctype html>
<html>
  <head>
    <title>P4B 2016</title>
  </head>
  <body>
    Programming For Biology 2016
  </body>
</html>
```
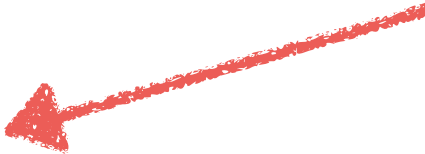
Doctype

Doctype

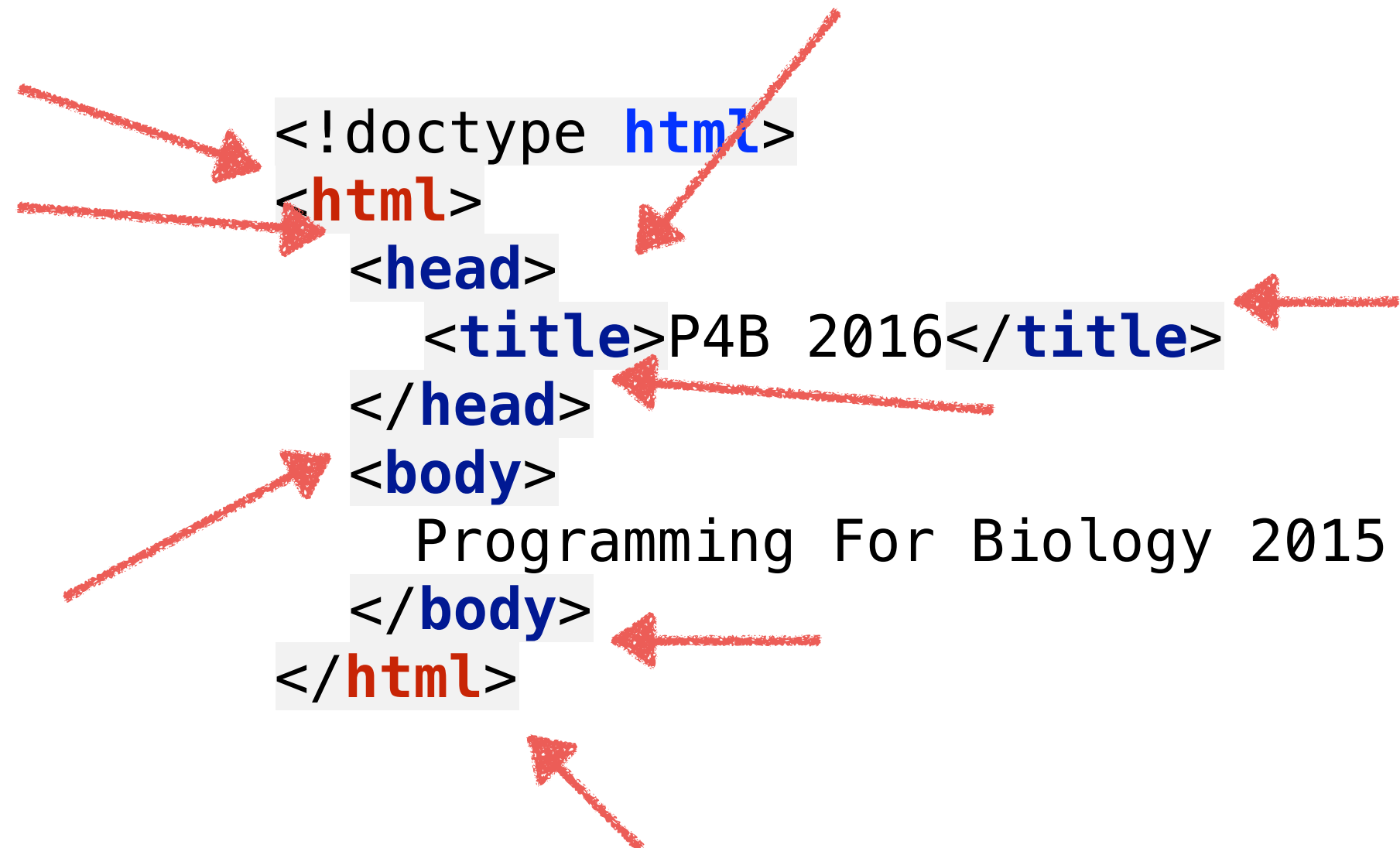```
<!doctype html>
<html>
   <head>
     <title>P4B 2016</title>
   </head>
   <body>
     Programming For Biology 2016
   </body>
</html>
```

The doctype declaration is used by web browsers

It declares that this file aspires to be correct HTML

# Elements

```
<!doctype html>
<html>
  <head>
    <title>P4B 2016</title>
  </head>
  <body>
    Programming For Biology 2015
  </body>
</html>
```

They provide **context** for their contents

Anatomy of an element

Start tag `<html>`

Angle bracket

Tag name

Angle bracket

Forward slash

must match

End tag `</html>`
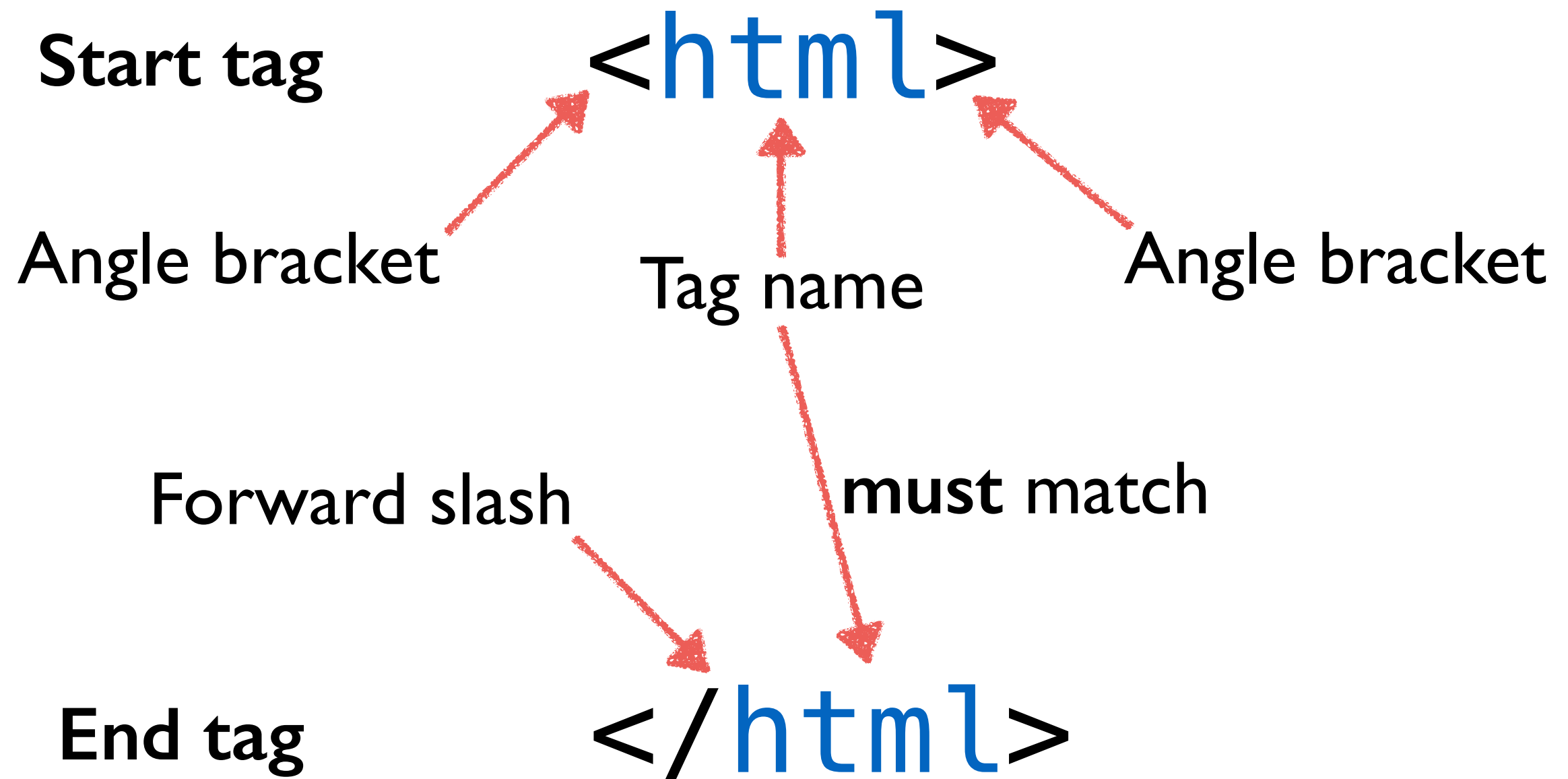
Elements

```
<!doctype html>
<html>
    <head>
        <title>P4B 2016</title>
    </head>
    <body>
        Programming For
        Biology 2016
    </body>
</html>
```

They **nest** and **balance** (like {braces} in perl)

A basic page

```html
<!doctype html>
<html>
  <head>
    <title>P4B 2016</title>
  </head>
  <body>
    Programming For
    Biology 2016
  </body>
</html>
```

All pages have <html>, <head> and <body>

`<html>`   Surrounds the entire document
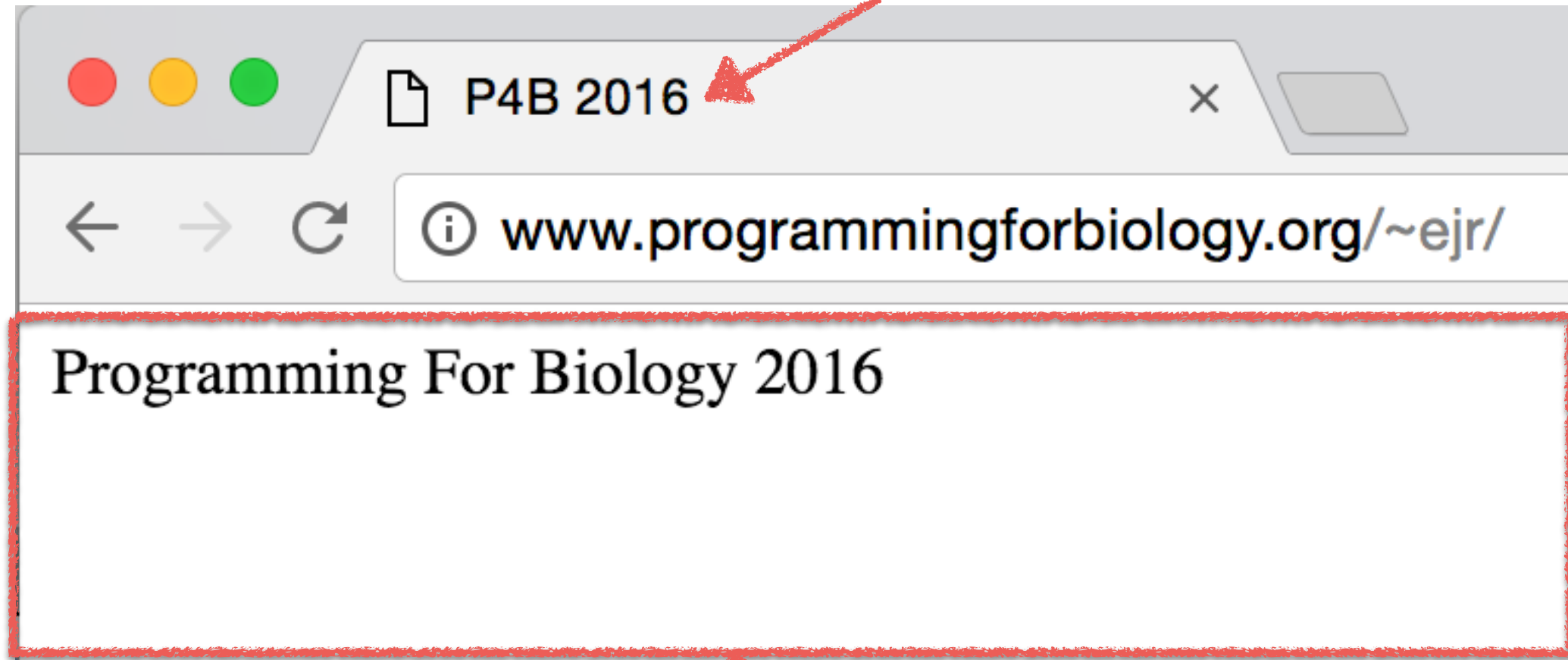
(except the doctype)

`<head>`   Contains metadata about the page

e.g. the page `<title>`

`<body>`   The part of the page that is displayed

A basic page

Page <title> (from <head>)

P4B 2016

www.programmingforbiology.org/~ejr/

Programming For Biology 2016

Contents of <body>

Text elements

Heading tag

This is a top level heading

`<h1>`This is a top level heading`</h1>`
`<h6>`This is the bottom level heading`</h6>`

This is the bottom level heading

Paragraph tag

`<p>`This is definitely a paragraph`</p>`

This is definitely a paragraph

Emphasis and Strong

That's `<em>`*exactly*`</em>` what I mean – I am `<strong>`**sick**`</strong>` of this slide

That's *exactly* what I mean - I am **sick** of this slide

Comment Tag

`<!-- This is a comment. You won't see this on the web-->`

Lists

Two kinds:-

`<ol>`   Ordered list;   numbered items

`<ul>`   Unordered list;   bulleted items

Each may only contain:-

`<li>`   A list item

Lists

```html
<ol>
    <li>
        <h4>First things first</h4>
        <ul>
            <li>Who you know</li>
        </ul>
    </li>
    <li>
        <h4>Not</h4>
        <ul>
            <li>What you know</li>
            <li>What you can do with it</li>
        </ul>
    </li>
</ol>
```
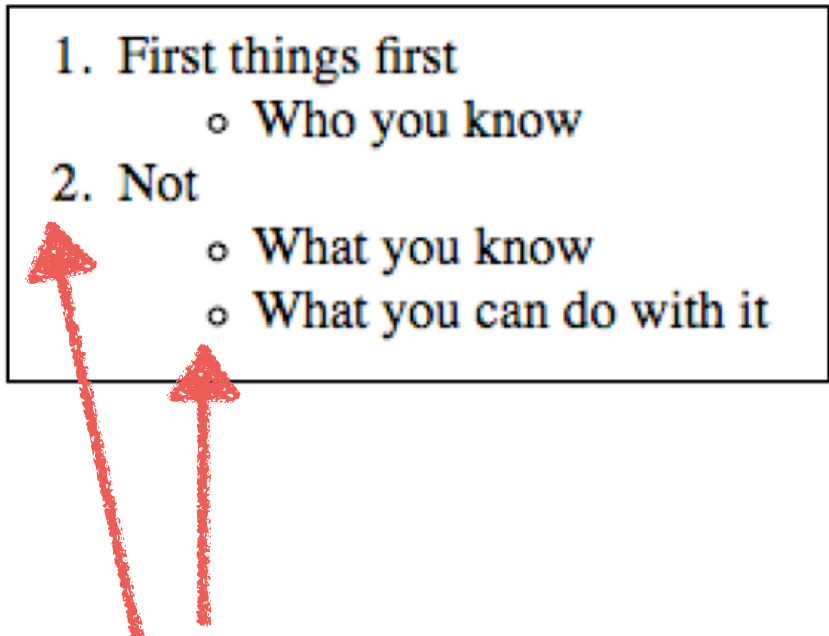
Lists may be **nested**

output:

1. First things first
   - Who you know
2. Not
   - What you know
   - What you can do with it

Numbers, bullets and indentation added by web browser

Link and attributes

Attributes describe an element but are not displayed

`<a href="http://nytimes.com">`

Attribute Name   Attribute Value

Links

"HTML Anchor element"

Absolute `<a href="http://nytimes.com">The New York Times</a>`

[The New York Times](#)

Relative `<a href="another_page.html">On The Same Site</a>`

Internal `<a href="#faqs">On The Same Page</a>`

↓

`<h2 id="faqs">FAQs</h2>`

Images

img   Image tag name

src   URL of image file

alt   Text description of image

```
<img src="images/Early.png" alt="An Appalachian Mud Squid">
```

Tables

`<table>`   This surrounds the entire table

`<tr>`   A table row

Every row in the table needs one of these

A table row contains cells.

Two cell tags:-

`<th>`   Header item

`<td>`   Data item

Each row should have the same number of th or td elements!

Tables

```html
<table>

    <tr>
        <th>ID</th>
        <th>Length</th>
        <th>Sequence</th>
    </tr>

    <tr>
        <td>M12277.1</td>
        <td>22</td>
        <td>CAACCCCCCGCCCTCAGATCCC</td>
    </tr>

    <tr>
        <td>X83548.1</td>
        <td>21</td>
        <td>CAAAGGAGGCGCTAAGCGCCA</td>
    </tr>

</table>
```
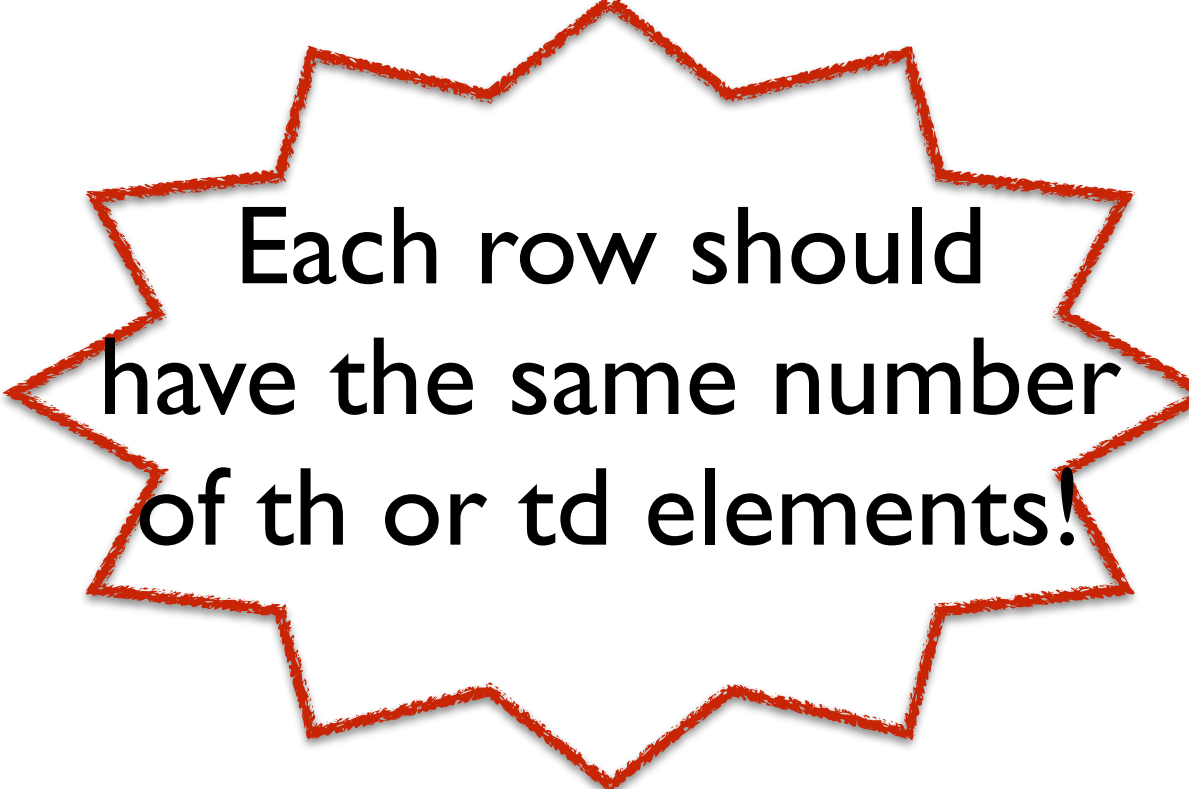
ID	Length	Sequence
M12277.1	61	CAACCCCCCGCCCTCAGATCCCTAATCCCATCGCAAGCATCAGACTCCCTCCAAAGCAGGC
X83548.1	21	CAAAGGAGGCGCTAAGCGCCA

Forms

```
<form name="yourname" action="handleMyForm.pl">
```



```
</form>
```

Forms allow browsers to send data entered by a user to the server

The **action** attribute specifies the URL that will receive the data

Forms

```
<form name="yourname" action="handleMyForm.pl">

    <input type="text" name="firstname" />


    <input type="text" name="lastname" />


</form>
```

output:

| What | Is This? |

Forms

```html
<form name="yourname" action="handleMyForm.pl">
    <label for="fname">First name:</label>
    <input id="fname" type="text" name="firstname" />

    <label for="lname">Last name:</label>
    <input id="lname" type="text" name="lastname" />

</form>
```

output:

First name: [Joe]   Last name: [Mulvaney]

Forms

```
<form name="yourname" action="handleMyForm.pl">
    <label for="fname">First name:</label>
    <input id="fname" type="text" name="firstname" />

    <label for="lname">Last name:</label>
    <input id="lname" type="text" name="lastname" />

    <input type="submit" value="Submit" />
</form>
```

output:

First name: [Joe] Last name: [Mulvaney] [Submit]

(There are many other **types** of <**input**>)

Forms

```html
<form name="yourname" action="handleMyForm.pl">
    <label for="fname">First name:</label>
    <input id="fname" type="text" name="firstname" />

    <label for="lname">Last name:</label>
    <input id="lname" type="text" name="lastname" />

    <input type="submit" value="Submit" />
</form>
```

output:

First name: [Joe   ]   Last name: [Mulvaney   ]   [Submit]

1. Package up every **<input>** with a **name** attribute

2. Send it to the URL in the **<form>**'s **action**

Organizing content with tags

- Sometimes you need to organize content but a table or list is not appropriate

`<div>`   Define a **block** of content

`<span>`   Mark a piece of content **inline**

`<section>`   "Semantic tags": Define a block of
`<aside>`   content more descriptively.
`<footer>`

`<nav>`   Recently introduced; not in IE8

Organizing content with attributes

id   Must be unique on the page

class   Groups related information
              A tag can have many classes
              separated by spaces

These are the most important,
most commonly used attributes

Organized Content

```
<table>
    <tr>
        <th>ID</th>
        <th>Length</th>
        <th>Sequence</th>
    </tr>

    <tr>
        <td>M12277.1</td>
        <td>22</td>
        <td>CAACCCCCCGCCCTCAGATCCC</td>
    </tr>

    <tr>
        <td>X83548.1</td>
        <td>21</td>
        <td>CAAAGGAGGCGCTAAGCGCCA</td>
    </tr>
</table>
```

Organized Content

?

output:

Results

ID	Length	Sequence
M12277.1	61	CAACCCCCGGATCCTCAGATCCC
X83548.1	21	CAAAGGAGGGATCCAAGCGCCA

```html
<table>
    <tr>
        <th>ID</th>
        <th>Length</th>
        <th>Sequence</th>
    </tr>

    <tr>
        <td>M12277.1</td>
        <td>61</td>
        <td class="sequence">CAACCCCC<span class="motif">GGATCC</span>TCAGATCCC</td>
    </tr>

    <tr>
        <td>X83548.1</td>
        <td>21</td>
        <td class="sequence">CAAAGGAG<span class="motif">GGATCC</span>AAGCGCCA</td>
    </tr>
</table>
```

HTML + CSS = Web

```html
<body>
<div id="wrap">
    <div id="header"><h1>Simple 2 column CSS layout, final layout</h1></div>
    <div id="nav">
        <ul>
            <li><a href="#">Option 1</a></li>
            <li><a href="#">Option 2</a></li>
            <li><a href="#">Option 3</a></li>

            <li><a href="#">Option 4</a></li>
            <li><a href="#">Option 5</a></li>
        </ul>
    </div>
    <div id="main">
        <h2>Column 1</h2>
        <p><a href="/">456 Berea Street Home</a></p>

        <p><a href="/lab/developing_with_web_standards/csslayout/2-col/">Simple 2 column CSS layout<
        <p>Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Mauris vel magna. Mauris risus
        <p>Nulla a lacus. Nulla facilisi. Lorem ipsum dolor sit amet, consectetuer adipiscing elit.
        <p>Aenean tempor. Mauris tortor quam, elementum eu, convallis a, semper quis, purus. Cras at
        <h3>Consectetuer adipiscing elit</h3>
        <p>Nulla dictum.                                                 malesuada sed, ligula. Sed

        <p>Maecenas eu v                                                 vamus pellentesque, lacus eu
    </div>
    <div id="sidebar">
        <h2>Column 2</h2>
        <p>Lorem ipsum d                                                 ris vel magna.</p>
        <ul>
            <li><a href="#">Link 1</a></li>

            <li><a href="#">Link 2</a></li>
            <li><a href="#">Link 3</a></li>
            <li><a href="#">Link 4</a></li>
            <li><a href="#">Link 5</a></li>
            <li><a href="#">Link 6</a></li>
            <li><a href="#">Link 7</a></li>

            <li><a href="#">Link 8</a></li>
            <li><a href="#">Link 9</a></li>
            <li><a href="#">Link 10</a></li>
            <li><a href="#">Link 11</a></li>
            <li><a href="#">Link 12</a></li>
            <li><a href="#">Link 13</a></li>

            <li><a href="#">Link 14</a></li>
            <li><a href="#">Link 15</a></li>
        </ul>
    </div>
    <div id="footer">
        <p>Footer</p>
    </div>

</div>
</body>
```

Content

+

```css
<style type="text/css">
body,html {
    margin:0;
    padding:0;
    color:#000;
    background:#a7a09a;
}
#wrap {
    width:750px;
    margin:0 auto;
    background:#99c;
}
#header {
    padding:5px 10px;
    background:#ddd;
}
h1 {
    margin:0;
}
#nav {
    padding:5px 10px;
    background:                9;
}
#                         m
#                         m

                          e;
#

#
#main {
    float:left;
    width:480px;
    padding:10px;
    background:#9c9;
}
h2 {
    margin:0 0 1em;
}
#sidebar {
    float:right;
    width:230px;
    padding:10px;
    background:#99c;
}
#footer {
    clear:both;
    padding:5px 10px;
    background:#cc9;
}
#footer p {
    margin:0;
}
* html #footer {
    height:1px;
}
</style>
```

Style

=

Simple 2 column CSS layout, final layout

Option 1 Option 2 Option 3 Option 4 Option 5

Column 1

456 Berea Street Home

Simple 2 column CSS layout

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Mauris vel magna. Mauris risus nunc, tristique varius, gravida in, lacinia vel, elit. Nam ornare, felis non faucibus molestie, nulla augue adipiscing mauris, a nonummy diam ligula ut risus. Praesent varius. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus.

Nulla a lacus. Nulla facilisi. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Fusce pulvinar lobortis purus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Donec dapibus, magna quis pulvinar laoreet, dolor neque lacinia arcu, et luctus mi erat vestibulum sem. Mauris faucibus iaculis lacus. Aliquam nec ante in quam sollicitudin congue. Quisque congue egestas elit. Quisque viverra. Donec feugiat elementum est. Etiam nec lorem.

Aenean tempor. Mauris tortor quam, elementum eu, convallis a, semper quis, purus. Cras at tortor in purus tincidunt tristique. In hac habitasse platea dictumst. Ut eu lectus eu metus molestie iaculis. In ornare. Donec at enim vel erat tempor congue. Nullam varius. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Nulla feugiat hendrerit risus. Integer enim velit, gravida id, sollicitudin at, consequat sit amet, leo. Fusce imperdiet condimentum velit. Phasellus nonummy interdum est. Suspendisse quam.

Consectetuer adipiscing elit

Nulla dictum. Praesent turpis libero, pretium in, pretium ac, malesuada sed, ligula. Sed a urna eu ipsum luctus faucibus. Curabitur fringilla. Suspendisse sit amet quam. Sed vel pede id libero luctus fermentum. Vestibulum volutpat tempor lectus. Vivamus convallis tempus ante. Nullam adipiscing dui blandit ipsum. Nullam convallis lacus ut quam. Mauris semper elit at ante. Vivamus tristique. Pellentesque pharetra ante at pede. In ultrices arcu vitae purus. In rutrum, erat at mollis consequat, leo massa consequat libero, ac vestibulum libero tellus sed risus. Lorem ipsum dolor sit amet, consectetuer adipiscing elit.

Maecenas eu velit nec magna venenatis consequat. In neque. Vivamus pellentesque, lacus eu aliquet semper, lorem metus rhoncus metus, a ornare orci ante a diam. Nunc sem nisl, aliquet quis, elementum nec, imperdiet in, wisi. Proin in lorem. Etiam molestie diam eget ante. Morbi quis tortor id lacus mollis venenatis. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Aliquam vel orci sit amet tellus mollis eleifend. Donec urna diam, viverra eget, ultricies gravida, ultrices eu, erat. Proin vel arcu. Sed diam. Cras hendrerit arcu sed augue. Sed justo felis, luctus a, accumsan in, tincidunt facilisis, libero. Phasellus eu eros.

Column 2

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Mauris vel magna.

- Link 1
- Link 2
- Link 3
- Link 4
- Link 5
- Link 6
- Link 7
- Link 8
- Link 9
- Link 10
- Link 11
- Link 12
- Link 13
- Link 14
- Link 15

Footer

Cascading Style Sheets

- Help separate **content** from **appearance**

- You don't need to write your own!

  - Many excellent style sheets available to you

  - I like Twitter Bootstrap (getbootstrap.com)

How CSS works

- Statements consist of

  - Selectors

    - *Which* element(s)?

  - Declarations

    - *What* will you do to them?

A declaration

Selector

```css
body {
    background-color: orange;
}
```

CSS Example

```css
.sequence {
    font-family: monospace;
}

.motif {
    color: orange;
}
```

HTML

```
<table>
    <tr>
        <th>ID</th>
        <th>Length</th>
        <th>Sequence</th>
    </tr>

    <tr>
        <td>M12277.1</td>
        <td>61</td>
        <td class="sequence">CAACCCCC<span class="motif">GGATCC</span>TCAGATCCC</td>
    </tr>

    <tr>
        <td>X83548.1</td>
        <td>21</td>
        <td class="sequence">CAAAGGAG<span class="motif">GGATCC</span>AAGCGCCA</td>
    </tr>
</table>
```

CSS

```
.sequence {
    font-family: monospace;
}

.motif {
    color: orange;
}
```

+

Results

ID	Length	Sequence
M12277.1	61	CAACCCCC**GGATCC**TCAGATCCC
X83548.1	21	CAAAGGAG**GGATCC**AAGCGCCA

CSS Selectors

```css
body {

}
```
Match the "body" element

```css
.motif {

}
```
Match all elements
with class="motif"

```css
#firstname {

}
```
Match the element
with id="firstname"

DIY

- There is a sample web site in your home directory

    $ cd ~/public_html

- Take a look at index.html

- Then open your favorite browser and go to

    http://www.programmingforbiology.org/~<YOUR USER NAME>/

Scraping

- We can parse web pages like any other text

- Use `curl` or `wget` to download an html file

```
$ curl http://programmingforbiology.org | less

$ wget http://programmingforbiology.org/index.html
$ less index.html
```

```perl
#!/usr/bin/perl
# A quick script to parse PFAM search results
use strict;
use warnings;

my $search = shift;
system ("wget -q -O pfam_search.txt 'http://pfam.xfam.org/search/
keyword?query=" . $search . "'");

open (FILE, "<", "pfam_search.txt") or die "Cannot open file: $!
\n";

# create a hash of the ids in our web page
# We use a hash as an easy way to eliminate duplicates
my %pfam_ids;
while (my $line = <FILE>) {
    if ($line =~ /(PF\d+)/) {
        $pfam_ids{$1}++;
    }
}

# Print the IDs on one line
print join ("\t", $search, sort keys %pfam_ids), "\n";
```

`</html>`