Searching scientific literature
Reading scientific literature
Writing scientific literature

These slides were originally prepared by Dr. Deepak Barua at IISER Pune. They have been modified by various course instructors (including me) over time.

Searching scientific literature

What is ‘scientific literature’?

A collection of written work

Three basic levels of the scientific literature:

Primary
Original research written by the those who conducted the studies
For e.g., journal articles, conference proceedings, patents, dissertations

Secondary
Uses primary sources to review, synthesize, topics
For e.g., Reviews, textbooks, encyclopedias

Tertiary
Aid for searching primary and secondary literature
For e.g., indexes, directories, bibliographies

Why bother with primary literature?

Journal papers are current

Textbooks are often years out of date

You can get enough details to replicate what you read about

Adapt cutting edge ideas and techniques to your own research

Training of critical faculties

Because one day soon you could be writing papers too!

Format of scientific literature

Pre-Prints
Manuscripts that have yet to undergo peer review. The immediate distribution of preprints allows authors to receive early feedback from their peers, which may be helpful in revising and preparing articles for submission.

Conference proceedings
Proceedings are the collection of academic papers that are published in the context of an academic conference.

Abstracts and indexes

Journal articles

Bibliographies
A systematic list of books and other works such as journal article

Databases

Books

Textbooks/dictionaries/encyclopedias

Format of scientific literature

Depends on the journal but are usually in these formats

- Research report
- Editorials
- Reviews
- Opinions
- Perspectives/synthesis
- Technical notes/protocol & methodology articles

Format of scientific literature

LETTER

doi:10.1038/nature10570

Lowland–upland migration of sauropod dinosaurs during the Late Jurassic epoch

Henry C. Fricke[1], Justin Hencecroth[1] & Marie E. Hoerner[1]†

Sauropod dinosaurs were the largest vertebrates ever to walk the Earth, and as mega-herbivores they were important parts of terrestrial ecosystems. In the Late Jurassic-aged Morrison depositional basin of western North America, these animals occupied lowland river-floodplain settings characterized by a seasonally dry climate[1,2]. Massive herbivores with high nutritional and water needs could periodically experience nutritional and water stress under these conditions, and thus the common occurrence of sauropods in this basin has remained a paradox. Energetic arguments and mammalian analogues have been used to suggest that migration allowed sauropods access to food and water resources over a wide region or during times of drought or both[3,4], but there has been no direct support for these hypotheses. Here we compare oxygen isotope ratios ($\delta^{18}O$) of tooth-enamel carbonate from the sauropod *Camarasaurus* with those of ancient soil, lake and wetland (that is, 'authigenic') carbonates that formed in lowland settings. We demonstrate that certain populations of these animals did in fact undertake seasonal migrations of several hundred kilometres from lowland to upland environments. This ability to describe patterns of sauropod movement will help to elucidate the role that migration played in the ecology and evolution of gigantism of these and associated dinosaurs.

Inferring the behaviour of ancient organisms is difficult, but geochemical information preserved in their fossil remains can provide such an opportunity. This study of sauropod dinosaur behaviour relies on the fact that $\delta^{18}O$ values of surface waters ($\delta^{18}O_{sf}$; for example streams, lakes) vary significantly over any given landscape in response to differences in aridity and elevation among other environmental factors[5,6]. Authigenic carbonates ($CaCO_3$) form in basin soils, lakes and wetlands, and record the oxygen isotopic characteristics of these host isotopic domains when they precipitate. Similarly vertebrate tooth enamel (bioapatite $Ca_5(PO_4, CO_3)_3(OH, CO_3)$) records the oxygen isotope characteristics of the surface water reservoirs that serve as their drinking water[7,8]. If $\delta^{18}O_{sf}$ inferred from 'non-migratory' authigenic carbonates and from dinosaur tooth enamel differ, then it can be concluded that dinosaurs were drinking water that fell outside the basin and thus they travelled outside it.

To use this approach we analysed enamel carbonate from teeth ($n = 32$) of *Camarasaurus* sp. and *Camarasaurus lentus* collected at Thermopolis, Wyoming, and Dinosaur National Monument, Utah (DNM), respectively (Fig. 1a). Palaeosol and lacustrian carbonates were also analysed from DNM ($n = 38$; see Supplementary Information for details on methods and statistics). In addition, we used published $\delta^{18}O$ data obtained from a variety of authigenic carbonates found over the entire Morrison basin including the Thermopolis area[9–12]. Comparisons of isotopic data from co-occurring authigenic carbonates and tooth enamel, from tooth-enamel carbonate and tooth-enamel phosphate, and from single teeth indicate that primary palaeobiological information is preserved in tooth enamel (see Supplementary Information for more details about diagenesis).

To estimate $\delta^{18}O_{sf}$ using dinosaur tooth enamel, it is assumed that they fractionated oxygen isotopes in a manner similar to all water-dependent vertebrates studied so far, including birds, mammals and reptiles[7,8]. To estimate $\delta^{18}O_{sf}$ using authigenic carbonate, it is assumed that oxygen isotope fractionation occurred at 24 °C, a temperature consistent with modelled mean annual temperature for the region[13] (see

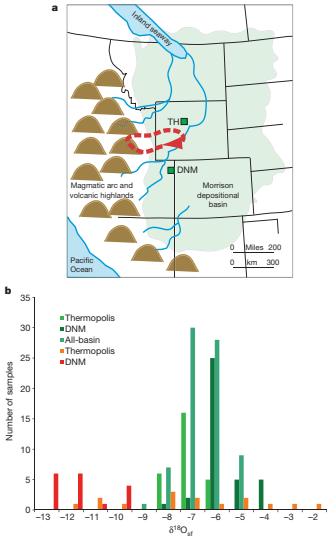

Figure 1 | Fossil localities, inferred oxygen isotope ratios of surface water and possible *Camarasaurus* migration routes. a, Palaeogeography of western North America during late Jurassic/Morrison time (after refs 2, 9), including fossil localities and one hypothetical migration route. **b**, $\delta^{18}O_{sf}$ estimated using tooth enamel (reds) and authigenic carbonates (greens; Thermopolis data from ref. 11; all-basin data from refs 9, 10, 12). See text and Supplementary Information for details.

[1]Department of Geology, Colorado College, Colorado Springs, Colorado 80903, USA. †Present address: Department of Geophysical Sciences, University of Chicago, Chicago, Illinois 60637, USA.

Short reports of original research
Focused on an outstanding finding whose importance means that it will be of interest to scientists in other fields.

ARTICLE

doi:10.1038/nature10716

DNA–binding factors shape the mouse methylome at distal regulatory regions

Michael B. Stadler[1,2]*, Rabih Murr[1]*, Lukas Burger[1,2]*, Robert Ivanek[1], Florian Lienert[1,3], Anne Schöler[1,2,3], Christiane Wirbelauer[1], Edward J. Oakeley[4], Dimos Gaidatzis[1,2], Vijay K. Tiwari[1] & Dirk Schübeler[1,3]

Methylation of cytosines is an essential epigenetic modification in mammalian genomes, yet the rules that govern methylation patterns remain largely elusive. To gain insights into this process, we generated base-pair-resolution mouse methylomes in stem cells and neuronal progenitors. Advanced quantitative analysis identified low-methylated regions (LMRs) with an average methylation of 30%. These represent CpG-poor distal regulatory regions as evidenced by location, DNase I hypersensitivity, presence of enhancer chromatin marks and enhancer activity in reporter assays. LMRs are occupied by DNA-binding factors and their binding is necessary and sufficient to create LMRs. A comparison of neuronal and stem-cell methylomes confirms this dependency, as cell-type-specific LMRs are occupied by cell-type-specific transcription factors. This study provides methylome references for the mouse and shows that DNA-binding factors locally influence DNA methylation, enabling the identification of active regulatory regions.

Methylation of cytosines in metazoan genomes adds epigenetic information onto DNA without changing the genetic information. It is thought to locally influence the regulatory potential of DNA, leading to changes in gene expression during development and disease[1,2]. Highly parallel sequencing enables to determine DNA methylation genome-wide at base-pair and single-molecule resolution in unprecedented breadth and detail[3,4]. Using whole-genome bisulphite sequencing (BisSeq), we generated such methylomes in mouse pluripotent embryonic stem (ES) cells and in neuronal progenitors (NP) and focused our analysis on methylation patterns that require high resolution for their detection. Using an analytical approach that quantifies DNA methylation locally, we identified a novel epigenome feature defined by localized reduced levels of methylation. These low-methylated regions (LMRs) show an average methylation of 30% and occur distal to promoters with little overlap with CpG islands. Contrasting LMRs with maps of histone modifications, DNase I hypersensitivity and several DNA-binding factors revealed that they are distal regulatory regions. Using the insulator protein CTCF and the transcriptional repressor REST as paradigms we show that DNA-binding factors are necessary and sufficient to create LMRs. LMRs form dynamically during differentiation driven by cell-type-specific factors, revealing a continuous crosstalk between DNA-binding factors and local DNA methylation.

Features of the mouse methylome

To obtain base-pair-resolution methylomes, we performed bisulphite conversion of DNA from mouse ES cells and corresponding NP derived through an established *in vitro* differentiation system[5–7]. High-coverage Illumina sequencing resulted in 77.3 billion mappable methylome bases. In parallel, we performed genomic sequencing to fully reconstruct the genotype together with available sequence data[8]. Our results show that cytosine methylation lies mostly in the context of CpG dinucleotides, and we observe non-CG methylation at levels that correspond to previous observations in human ES cells[3] (Supplementary Fig. 1). Similar to the human methylome, however, non-CG methylation is generally found in regions already containing CpG methylation, and because it appears as a rare event we focused our further analysis on CpG methylation. A global analysis yielded the anticipated DNA methylation pattern for mammalian cells[3,4,7,9]. CpG methylation is low at promoter regions, whereas genic and intergenic regions show high methylation levels (Supplementary Fig. 1). The methylation frequency of individual CpGs has a bimodal distribution: the majority of CpGs show a high percentage of methylation, whereas a smaller group of cytosines are unmethylated (Fig. 1a).

Identification of low-methylated regions

In addition to this expected binary behaviour of CpG methylation (either hypo- or hyper-methylated), we also noticed cytosines with intermediate, yet low levels of methylation in the range of 10–50% (Fig. 1a). More than 4% of all CpGs (almost 1 million in the mouse genome) show this characteristic. When plotting methylation along chromosomal regions, cytosines with such methylation are not randomly distributed but cluster locally (Fig. 1b). We thus developed a computational approach that segments methylomes into regions of differential methylation. The applied Hidden Markov Model (HMM) takes into account only DNA methylation and no genome information such as CpG density or functional annotations. This segmentation identified three distinct classes: fully methylated regions (FMRs), unmethylated regions (UMRs) and low-methylated regions (LMRs) (Fig. 1a–c). It correctly identifies unmethylated CpG islands as UMRs[3,4], whereas most of the genome is classified as FMR, consistent with the fact that the majority of CpGs are methylated. The segmentation further identified LMRs at positions where local dips are visible in the raw data profiles, further supporting that these have distinct and identifiable methylation patterns (Fig. 1a–c). Importantly, LMRs are generally not CpG islands as they have lower CpG content, are shorter and they further reside distal to transcriptional start sites (TSS; Fig. 1c–f and Supplementary Fig. 1).

LMRs are distal regulatory regions

Interestingly, LMRs show evolutionary conservation (Fig. 2a), suggesting that they are of functional relevance. Further, they contain DNase-I-hypersensitive sites (DHS), a unique chromatin state that

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland. [2]Swiss Institute of Bioinformatics, 4058 Basel, Switzerland. [3]Faculty of Science, University of Basel, 4056 Basel, Switzerland. [4]Novartis Institutes for BioMedical Research, Biomarker Development, 4056 Basel, Switzerland.
*These authors contributed equally to this work.

Original reports whose conclusions represent a substantial advance in understanding of an important problem and have immediate, far-reaching implications.

Life cycle of scientific information

The Idea – What? Why? How?

Research – Experiments, analysis, writing

Peer Review

Publication

Evaluation of the impact of the idea

"Publications are at once the end-product of scientific research and the raw material for future research." -- A.F Spillhaus

Life cycle of scientific information

Peer Review Process
Impartial scrutiny by experts

Why Peer review:
Division of labor
Diversity of opinion
Expertise in field

Other features of the review system:
How are reviewers selected?
Reviewer anonymity?
What criteria are used to review the work?

Problems:
Slow
Can be biased (possibility of abuse by reviewers)
Inconsistent

Alternatives? ……..double-blind reviews, open reviews

Three different kinds of reviews

1. Single blinded

The names of the reviewers are hidden from the author, but reviewers know the identity of the author(s). This is probably the most widely used method of reviewing and is, by far, the most common type.

Advantages
The anonymity of reviewer allows them to be honest and open without fear of criticism from the authors
Knowing the identity (and affiliation) of the author allows them to use prior knowledge to assist the article assessment

Disadvantages
Personality Influences: Reviewers who know authors with a dazzling record may be over-impressed by their work and not scrutinize the article carefully
Possible discrimination based on sexism and regionalism: Reviewers may be biased against certain individuals, or overly suspicious of research emerging from some regions of the world.

http://pkpschool.sfu.ca/becoming-an-editor/module-7/unit-1-the-different-types-of-peer-review/

Three different kinds of reviews

2. Double blinded

In double blinded reviewing the identity of both the reviewer and author is hidden from each other. The reviewers do not know who the authors are, and the authors do not know who the reviewers are.

Advantages
This method ensures fairness of judgement, without bias towards certain authors or regions of the world
Confidentiality of both author and reviewer may provide some protection against criticism for both the authors and reviewers

Disadvantage
This method is not foolproof as it may be easy for reviewers to discover the identities of the authors (either because of the area of research, the references or the writing style)
It may be argued that knowing the author's identity and affiliation will help the reviewer to make an informed judgement of the article – and not knowing this may disadvantage the review

Three different kinds of reviews

3. Open review

Open review does not attempt to conceal the identity of authors or reviewers. The identities of both reviewers and authors are known to each other.

Advantages
Open peer-review, and the transparency it brings to the process, encourages accountability, encourages civility, and generally improves the quality of the review (and ultimately the article)
Since an open review system may report the reviewers' names next to the accepted, published article, it will persuade reviewers to do a thorough job

Disadvantages
Some reviewers might be concerned about the consequences of being identified as the source of a negative review, and so may refuse to review for a journal using an open system
Reviewers might be reluctant to criticize the work of more senior researchers on whom they may depend for career advancement. This may be a particular problem in smaller research communities, and more prevalent in some regions of the world than others

How do you evaluate a web resource?

Authority

Is it clear who created the webpage, and can you contact them?

What are the author's credentials - educational background, past writings, or experience in this area?

What is the purpose of this site? Is there anything to suggest the information found there might not be objective or reliable?

Is the author affiliated with, or speaking on behalf of a government agency, an educational institution, or a respected non-profit organization? If so, the information found there might be more reliable than that found in the sites of commercial sites or Internet service providers.

Conflict of interest

How do you evaluate a web resource?

Credibility
Is the author clearly indicated?
Is there evidence of quality control?

Timeliness
What is the date of creation or latest revision of the site (if available)? A site that has not been updated recently may have dead links and the information found on the page may also be outdated.

Documentation
Where did the information included in the site come from? Are sources listed? Is there a bibliography included?

How do you evaluate a web resource?

Corroboration
It is *always* a good idea to avoid using information found in a web page without checking at least a few of the facts contained there against other sources.

Print Equivalents
Electronic sources that have print equivalents tend to be more reliable because they have undergone the same editorial review as their printed counterparts.

Databases

A collection of information that can be **searched**, and from which the individual pieces of information inside can be **retrieved**.

Print indexes:
Paper databases
Organized citations/abstracts of work in a series of books

Digital Databases:
Use digital tools to search, filter, and retrieve information much more quickly and efficiently

Bibliographic Databases:
Subject databases

Databases

General Search databases

Web of Science
Lexis-Nexis
Jstor
Google Scholar
Academic Search Complete
Emerald
…….etc

Specialized Search databases

PubMed
Chemical Abstracts Online
Scopus
Science Direct
…….etc

Databases

Each database will have its own strengths and weaknesses

Some are available/open to the public, while others require registration and/or subscriptions

Searching the scientific literature

Reasons for searching

Defining your topic

Narrowing or broadening your topic

Refining and focusing your search

Approaches to searching
Subject searches
Citation searches

Reasons for searching

General awareness
Browsing general science journals
Science news media
Database searches – content alerts

Practical everyday needs
Specific information
For e.g., method/technique, math equation, definitions

Exhaustive research
Need to identify ‘all’ relevant information in a field

Defining your topic

Understand your assignment

Select a topic that interests you!!

Select a topic that you have some basic understanding of

Select a topic that is unique
(..or takes a unique perspective of a topic)

Generate questions from your topic

Brainstorming
Free associations
Inventory building
Categorize/classify

Concept Mapping
Visual/graphic representation
Relationships between units

Creating a search strategy

Summarize topic
Clearly state the topic/research question in one or two statements
For e.g., how do giant squids find their prey in the absence of light?

Identify concepts
Identify the main concepts in your topic/research question
For e.g., how do giant squids find their prey in the absence of light?

Select keywords
Find other synonyms/variants (or other associated words)
For e.g., Giant squid – squid, cephalopod
Find prey – feeding behavior; hunting; food
Absence of light – dark, optical sensors

Focusing and refining your search

Start with the secondary literature
Reviews/monographs
Web resources
Encyclopedia
Textbooks (library catalogues/databases)

Searching the primary literature (use tertiary literature)
Find appropriate database/search engine

Creating a search strategy (subject searching)

Connect your keywords (concepts)
Use of Boolean operators – AND, OR, NOT

Use truncated/wild card searches
Use of ?, * or other wild card operators to expand search

Other tips
Narrowing your search based on:
Region
Organism
Type of reference
Time of publication
Other

Hierarchical relationships between concepts
English vs. American spellings

Creating a search strategy (citation searching)

Looking backwards
Finding relevant citations
Finding other major work on the topic

Looking forward
Finding current work that has cited these references

Reading scientific literature

Why do I need to read a scientific paper?

- For general interest or background information
- To find out exactly what the latest developments are in a field
- To seek evidence to support or refute your ideas
- To broaden your avenues of research
- To find out how a certain piece of research was done

Title

- Starting point
- Should briefly describe the study
- Should clarify what your expectations are of the paper
- Should help decide if the study is relevant to your work (thus, titles should get attention)

Keywords

- **Keywords** are a tool to help indexers and search engines find relevant papers. If database search engines can find your journal manuscript, readers will be able to find it too. This will increase the number of people reading your manuscript.

- However, to be effective, Keywords must be chosen carefully. They should:
- **Represent** the content of your manuscript
- Be **specific** to your field or sub-field

What kind of paper?

- Original research?
- Review, opinion, perspective?
- Peer-reviewed or by invitation?
- High-impact journal?
- Author’s reputation?

What kind of paper?

- Papers and journals are judged by their citation rates and impact factors.
- Also, need to ask is this a specialist journal or general journal?

Impact factor

- Measures the average number of citations per number of articles published by a journal

- Often used as a proxy of the **relative importance** of a journal **within its field**

$$\text{Impact factor} = \frac{\text{\# citations in previous 2 years}}{\text{\# publications in previous 2 years}}$$

(NOTE: number of publications = number of ‘citable’ articles published)

Impact factor

- Highly discipline dependent
  Average time after publication before citation
  Number of citations per article

MOLECULAR BIOLOGY

rank	journal	total cites	Impact Factor	immediacy index	Cited half life
1	ANNU REV BIOCHEM	16889	30.016	3.677	9.7
2	CELL	142064	31.253	6.126	8.8
3	NAT MED	48632	27.553	5.469	6.1

ECOLOGY

rank	journal	total cites	Impact Factor	immediacy index	Cited half life
1	B AM MUS NAT HIST	2003	16.692	1.444	>10.0
2	TRENDS ECOL EVOL	16830	11.904	1.913	7.4
3	ANNU REV ECOL EVOL S	10653	10.161	0.133	>10.0

MEDICINE – Highest impact factors - >50 !!

- Number of citations vary depending on type of article
- Process not transparent – independent audit (Rossner et al. 2007) was not able to reproduce the rankings

Impact factor

Alternative to impact factors

- **Immediacy index**
  The number of citations the articles in a journal receive each year divided by the number of articles published
- **Cited half-life**
  Median age of the articles that were cited in *Journal Citation Reports* each year

- **Aggregate impact factors**
  Considers all citations in a field and total number of articles in field

- **h-index**

Impact factors have a large, but controversial, influence on how scientific research is perceived and evaluated

Impact factor

- Not intended (*but often used*) for evaluating individual researchers

- Prone to misuse
  Self citation (by author and journal)
  Review articles inflate IF

Organization of a scientific paper

- IMRAD
  Introduction, Methods, Results and Discussion

- Plus
  Title, abstract, authors, acknowledgements, declarations, references
  Tables and figures; legends

Organization of a scientific paper

Variations

Length versus **accessibility** to non-expert
Combined Results and Discussion
Methods at end
On-line supplements

Reading a scientific paper

- **This is not a novel!**
- No need for a **linear** approach
- Usual workflow among experts and reviewers
  1. Title
  2. Abstract
  3. Figures, tables
  3. Introduction, results
  4. Discussion
  5. Methods

Reading a scientific paper

- Why you are reading determines how you should read
- The abstract & introduction should tell you whether it is worth reading in depth or only worth skimming
- The answer will depend on what you are looking for

Reading a scientific paper

Just because its published, does not mean its right
Get used to doing peer review

- Struggle with the paper
- Be an active and not passive reader
- If at first you don’t understand, read and re-read, spiraling in on central points
- Get into questioning mode
- Doubt everything
- Nit-pick and find fault
- Move beyond the text
- Talk to other people
- Read commentaries
- Refer to papers cited
- Consult, dictionaries, textbooks, online links to references, figure legends to clarify things you do not understand

Blame the authors if…

- **Logical connections left out**
  Instead of saying why something was done, the procedure is simply described
- **Cluttered with jargon, acronyms, technical terms**
- **Lack of a clear road-map through the paper**
  side issues given equal air-time with main thread
- **Difficulties determining what was done**
  Ambiguous or sketchy description
- **Data mixed up with interpretation and speculation**

Critical assessment

- **Read the experimental results**
  Figures and tables can be evaluated first

- **Avoid reading the discussion section before results**
  Readers should evaluate results before reading the authors’ conclusions

- **Use your own judgment**

Critical assessment

- What **major/general** question/s does the paper address?
- What are the main conclusions of the paper?
- What evidence supports those conclusions?
- Do the data actually support the conclusions?
- What is the quality of the evidence?
- Why are the conclusions important?

What questions does the paper raise?

- **Descriptive research**

  often in early stages of our understanding can't formulate hypotheses until we know what is there. for e.g., DNA sequencing and microarray

- **Comparative research**

  ask how general or specific a phenomenon is

What questions does the paper raise?

- **Analytical or hypothesis-driven research**
  Test hypotheses - for e.g., mutations and their effects on functions

- **Methodological research**
  Find out new and better ways of doing things
  Describe new resources - for e.g., description of new homology search method, genome database

- **Significant papers combine all the above**

What are the main conclusions?

- Do they matter?
- Are they of general relevance?
- Are they broad in scope?
- Are they detailed but with far-reaching conclusions?
- Are they accessible to general audience?

Places to find information about the subject of the paper

- The title
- The abstract, and
- The introduction

Note: The discussion contains further ideas, but it is not worth reading the discussion in any detail until we have good idea what is being discussed.

Abstract and introduction

- Abstract should give you a brief summary of the paper's main finding
- Introduction provides a background to the paper and a rationale for the investigation in more detail than is possible
- Introduction gives a 'state-of-the-art' of our current understanding of the field
- The abstract and introduction help you to decide whether, why and how to read

Why it is a good idea to read introductions?

- They give you some idea of background information you need before starting the paper
- They give you an insight into the authors' starting point and approach to the subject

Why it is a good idea to read introductions?

- The abstract and introduction should explain why the paper was written
- They do not give detailed information, but should help you decide how much time to spend on the paper
- Introductory sections are an entry into a paper – never substitute for reading it properly

What evidence supports them?

- Look at results section and relevant tables and figures.
- May be one primary experiment to support a point.
- More often several different experiments or approaches combine to support a particular conclusion.
- First experiment might have several possible interpretations, and the later ones are designed to distinguish among these
- In the ideal case, the Discussion begins with a section of the form “These lines of evidence provide support for the conclusion that...."

Judging the quality of evidence

- You need to understand the methods thoroughly (may need to consult textbooks)
- You need to know the limits of the methods
- Separate fact from interpretation
- Are the results expected?
- Extraordinary claims require extraordinary evidence

Judging methods

- There must be a logical reason why the method can or may answer the question
- Defined and reproducible protocols must be followed
- Controls must be in place in order to rule out extraneous influences on the results

Judging the quality of evidence

- Look at details
- Assess them for plausibility
- The veracity of whole depends on the veracity of its parts!
  For e.g., look at gene lists, what is missing but expected, what is present, but unexpected?
- Where are the controls?

Why is it a good idea to read methods and materials?

- To know how it was done in order to understand what it means
- If you want to replicate an experiment, the methods section is indispensable
- To find stimulating ideas and make connections between different areas
- To adapt methodological approaches to our own experiments

Do the data support the conclusions?

- Data may be believable but does not support the conclusion the authors wish to reach
- Logical connection between the data and the interpretation is not sound (often hidden by bad writing)
- Might be other interpretations that are consistent with the data

Do the data support the conclusions?

- **Rule of thumb**
  If multiple approaches, multiple lines of evidence, from different directions, supporting the conclusions, then more credible.

- **Question assumptions!**
  Identify any implicit or hidden assumptions used by the authors in interpreting their data?

Titles

- Starting point
- Should briefly describe the study
- Clarify what your expectations are of the paper
- Decide if the study is relevant to your work (thus, title should get your attention)

Titles

- Effects of flowering time genes on seed germination
- Do flowering time genes affect seed germination in *Arabidopsis thaliana?*
- Effects of the major flowering time gene, *FLOWERING LOCUS C*, on seed germination in *Arabidopsis thaliana*
- The major flowering time gene, *FLOWERING LOCUS C*, regulates seed germination in *Arabidopsis thaliana*
- Characterization of affects of the major flowering time gene, *FLOWERING LOCUS C,* on seed germination at lower temperatures in naturally occurring populations of *Arabidopsis thaliana*

More specific ← → More universal

Abstract

- Abstracts should concisely summarize the study
- Should include

  WHAT? What was the main objective and the major conclusions of the study

  HOW? Briefly describe the methodology used

  WHY? What is the relevance/significance of the study

- Abstract lengths can vary (usually imposed by the journal) from 100-300 words

Now you should know what the paper is about, and this should help you decide if and how to read the rest of the paper

Introduction

Sets the context:
Historical - how has this field progressed
Scientific - what is the current understanding in this field
(accepted state of knowledge in a specialized field)

Outlines hypothesis; assumptions; predictions; experimental design; etc.

Set the theoretical framework for the paper

Materials and methods

- Details of methods and material used – so that another researcher can reproduce your study
- Rationale for using the methods used
- Limitations of the methods used
- Other details: sample sizes; controls; statistical analysis etc.

Results

Figures, tables, text

Should present what the data show – not what the interpretation of that data is

Should use appropriate legends that describe the figures and tables

Figures and tables should be self-explanatory
the reader should not have to go read the text to understand it.

Figures and tables are the ‘heart’ of the paper

Discussion/conclusions

- Interpretations of the results and the conclusions that are drawn from those interpretations
- Relate back to the original hypothesis/predictions/objectives from the introduction
- Place the results/interpretations/conclusions in the context of the ‘current understanding’ in the field
- Present alternative explanations/interpretations of the results, avenues for further research, etc.

Writing scientific literature

General flow

- Understand your assignment
- Know your audience
- Research relevant information
- Write an outline
  Organize your ideas
  Logical flow
- Write an introduction, body and conclusion
- Connect paragraphs to each other in a logical order
- Edit and proofread

Plagiarism

What is Plagiarism?
Plagiarism is the use of another person's words or ideas as if they were your own.

Can it be avoided?
Plagiarism can be avoided by using a recognized referencing style, and by paraphrasing or quoting correctly

Paraphrasing
Involves using someone else's ideas but expressing them in your own words. The emphasis is on expressing the ***idea in your own way to***

Quoting
Involves using someone else's words *exactly as they appear in their work and is clearly* identified by the use of quotation marks.

Paraphrasing

Until you fully understand the idea that you are interested in using, you are not ready to attempt to incorporate it into your work.

If you understand it, then you should be better able to express it in your own way.

The sources of paraphrased ideas need to be acknowledged using an appropriate referencing system and style.

Common question
How much do I need to change the original wording for it to be considered "acceptable paraphrasing"?'.

Some easy ways to identify when you have not paraphrased correctly is when,

- Style of writing suddenly changes
- Easily recognizable words or phrases.
- Google search gives the original source as a hit

Referencing

- Acknowledge your sources
- Allow the reader to verify the data/information
- Allow the reader to consult your sources independently
- Show the reader the depth and breath of your reading

- The two basic components of a **referencing system** are:
- The citation in the text of your writing to the source of the information
- The reference list in alphabetical order

Writing a literature review

What is a literature review?

A literature review is an **evaluative comparison of various pieces of research**. It is not just a set of summaries or a descriptive list of material. It shows the reader **what previous research has been done** in the field, **critiques previous methodology**, and **evaluates prior studies to show an information gap** which your own or others' research will/can fill. The information which follows is particularly relevant to a thesis literature review but can be applied to shorter reviews and thesis proposals.

Stages

- Select the topic
- Set the topic in context
- Search for information sources
  define the information you need to have; state it as a question; break it into components; identify concepts, keywords, synonyms; select an information source that matches the information need.
- Evaluate your information sources
  read the abstract; note everything which *may be* important
- Evaluate the search process
  have you got too little information? maybe you need to broaden the scope of your search, try different types of sources or explore other disciplines.
- Have you got too much information?

Organizing information

- Extract the information from your sources

- Organize the information

  Making a summary table will help you see common threads in your literature; keep notes about the source of the information; use a reference manager, for e.g., Mendeley

  Group your literature in the review, for e.g., group related studies together, review briefly any weaker studies or studies that share similar methods, devote more attention to groundbreaking, stronger studies

  Group studies by findings or methodology or theory

General guidelines

- Set up a framework for your research with clearly set goals stated in the introduction and summarized in the conclusion
- Show your reader that you have a clear understanding of the key concepts/ideas/studies/models related to your topic
- Talk about the history of your research area and any related controversies
- Discuss these ideas in a context appropriate for your own investigation
- Evaluate the work of others
- Clarify important definitions and terminology

General guidelines

- Develop the research space you will also indicate in the introduction and abstract
- Narrow the problem down; make the study feasible
- Structure the review, using headings as appropriate
- Use good expression and punctuation
- Use your referencing system correctly
- Use current literature as well as older sources
- Identify the range of resources you have used
- Write in an interesting way

Citations

- Citations recognize and acknowledge the intellectual property rights of authors - they are a matter of ethics and a defense against plagiarism
- Citations are used to show respect to previous scholars.
- Citations operate based on a mutual reward system - writers ‘pay’ other authors in citations
- Citations are tools of persuasion - writers use citations to give their statements greater authority
- Citations demonstrate familiarity with the field

Common criticisms

- Landmark studies are not included
- Outdated material is given too much emphasis
- Recent literature is not included
- The perspective is not wide enough
- The review is not sufficiently analytical
- The writer can’t discriminate between relevant and irrelevant material
- There is no coherence
- The literature is not related to the research question or hypotheses
- Sources are not correctly interpreted