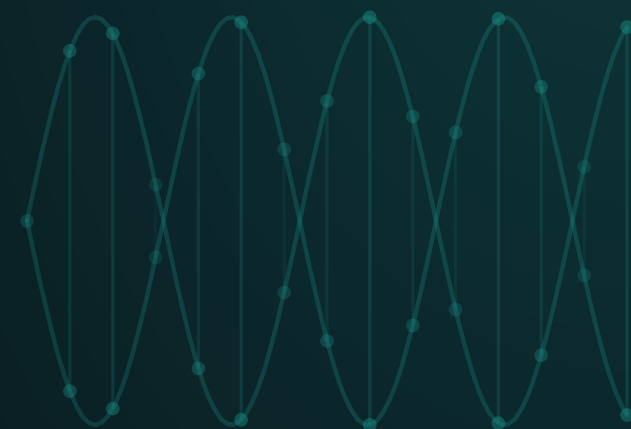


Lesson: when genetic findings don't apply to other populations



Genetics

Statistics

Social Studies

Health

Age 16+

Intermediate

A free lesson on reading data critically, using this site as the dataset. The class computes the same prediction in five populations and measures how wrong it gets — six points for one variant, fifty-two for another. Then we publish our own audit: 72.2% of the variants here were discovered in entirely European samples.

Free to read, use and adapt. No account, nothing to buy, nothing collected from anyone.

Full lesson, figures and updates:

<https://www.mygenelog.com/updates/lesson-does-this-number-travel>

What this is. A ready-to-run lesson on reading data critically, built on real genetics and on this site's own audit of itself. Free to use and adapt in any school, college or university. Nothing to buy from us, no account, no data collected from anyone.

- Level: from about age 16 through undergraduate. Assumes the words gene, allele and genotype, and the ability to square a decimal.
- Time: 50–60 minutes. The calculation takes about fifteen; the argument at the end can take the rest of the lesson and often should.
- You need: a calculator or a phone, and something to write on. The numbers are all printed below, so the lesson works with no internet in the room.
- It covers: allele frequency, carrier share and why the arithmetic depends on the biology, population structure, what a reference range is and how one goes wrong, sampling bias in a real research field, and auditing a dataset instead of trusting it.

Start here: a normal result, called a disease

There is a blood test almost everyone has had: a count of neutrophils, the white cells that deal with bacterial infection. If your count comes back below the reference range, that is neutropenia, and it is taken seriously. It can mean a drug needs

stopping, a job in healthcare needs pausing, a chemotherapy dose needs reducing, or that a bone marrow biopsy is warranted.

Now the problem. Around two-thirds of Black Americans, and the great majority of people from West, Central and East Africa, carry a variant that lowers the neutrophil count in the blood without making anybody ill. It is called Duffy-null, it is one of the commonest variants in the world, and it became common because it protects against a form of malaria.

For decades the reference range those counts were compared against was built largely on people who do not carry it. Healthy people were told they had a low white cell count. Some had treatment changed or withheld over it.

Nobody made an error of reasoning. The range was applied exactly as intended. It just did not travel — and the next hour is about how to check whether a number travels before it is used on someone.

Do this

Every number below is real, from published allele frequencies. Nobody in the room is tested, counted or asked anything about themselves.

1. Take a prediction first. “A study in one country finds that 25% of people carry a certain variant. If we go and measure a different country, how far off could that 25% be? A couple of points? Ten? More?” Write the guesses on the board.
2. Give them the tool. If an allele has frequency q , the share of people carrying at least one copy is $1 - (1 - q)^2$. Work through one together: at $q = 0.131$, that is $1 - 0.869^2 = 1 - 0.755 = 24.5\%$.
3. Start with the control, and say that it is the control. Hand out the frequencies for rs2075650: European 0.131, East Asian 0.097, African 0.131, South Asian 0.125, American 0.105. Each group computes the carrier share for each population. The answers land between 18.5% and 24.5% — a worst case of six points. A prediction made in Europe would be roughly right anywhere. This step is not filler. Without it the class will conclude that findings never transfer, which is false and would be a worse lesson than the one they came in with.
4. Now the second variant. rs9923231, in VKORC1 — the gene for the enzyme that warfarin blocks. Frequencies: European 0.388, East Asian 0.885, African 0.054, South Asian 0.145, American 0.411. Same calculation.
5. Then measure the error, which is the actual exercise. Take the European answer and use it as the prediction for each of the others. Subtract. How many percentage points wrong is it each time? Put the worst one on the board next to the guesses from step 1.
6. Finish with the one from the story. rs2814778, Duffy-null: European 0.006, East Asian 0.000, African 0.964, South Asian 0.000, American 0.078. Ask what a European-derived

expectation would say about an African population here.

7. One extra step for the last variant, and it is a good one. Duffy-null only changes the blood count in people carrying two copies. So the right calculation is q^2 , not $1 - (1 - q)^2$. Compute it: $0.964^2 = 92.9\%$ in African samples, $0.006^2 = 0.004\%$ in European ones. Ask why the formula changed. The answer is that it changed because the biology changed, not the maths — and a student who can say that has got the most transferable idea in the lesson.

Do not turn the class into the dataset. This lesson is about populations, and the obvious wrong move is to sort the room by background and count. Do not ask students where their family is from, do not group them by it, and do not use the class as an example of any population. A room is not a sample, and a student's ancestry is not a teaching aid.

As with every lesson here: collect no genetic or health data, from anyone, by any route. The numbers in this lesson are published population frequencies and they are all you need.

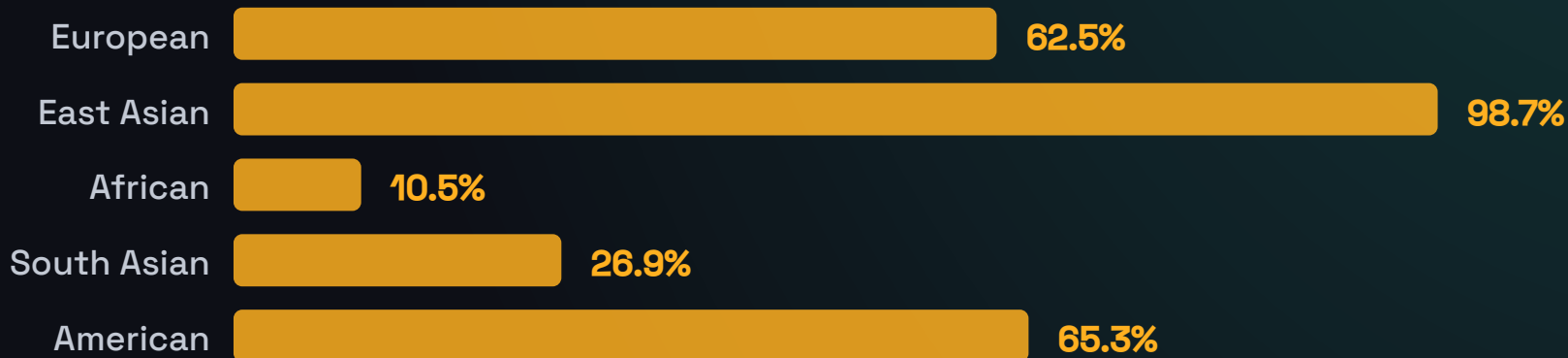
What the class just found

share of people carrying at least one copy — calculated from published allele frequencies

rs2075650 — the control: this one travels



rs9923231 — VKORC1, the warfarin target: this one does not



rs2814778 — Duffy-null: this one is almost only in one place



1000 Genomes phase 3 via Ensembl — carrier share is $1 - (1 - q)^2$, worked out here rather than quoted

Three variants, one calculation, three completely different answers about whether a number found in one place can be used in another. The top one is the reason this is a question and not an accusation: sometimes it transfers fine.

The middle variant is the useful one to sit with. A prediction made in Europe would be 36 points too low for an East Asian population and 52 points too high for an African one — from the same correct arithmetic, applied to the wrong people.

And this is not an abstract variant. It sits in the gene for warfarin’s target, and warfarin is a drug where the gap between too little and too much is small enough that people are tested repeatedly while their dose is found.

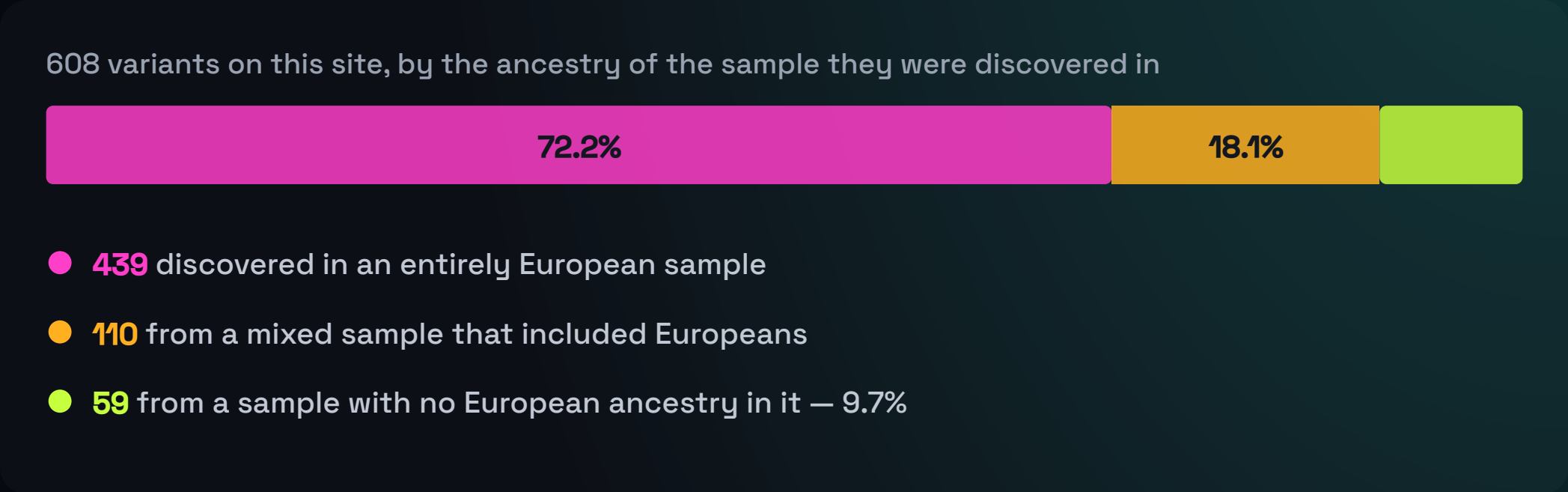
Real data: we audited ourselves for this lesson

It would be easy to teach this as something other people got wrong. So we ran the same question at our own catalogue.

Every variant on this site names the paper it came from. We took all 187 of those papers, asked the GWAS Catalog which ancestries each study’s discovery sample actually contained, and counted.

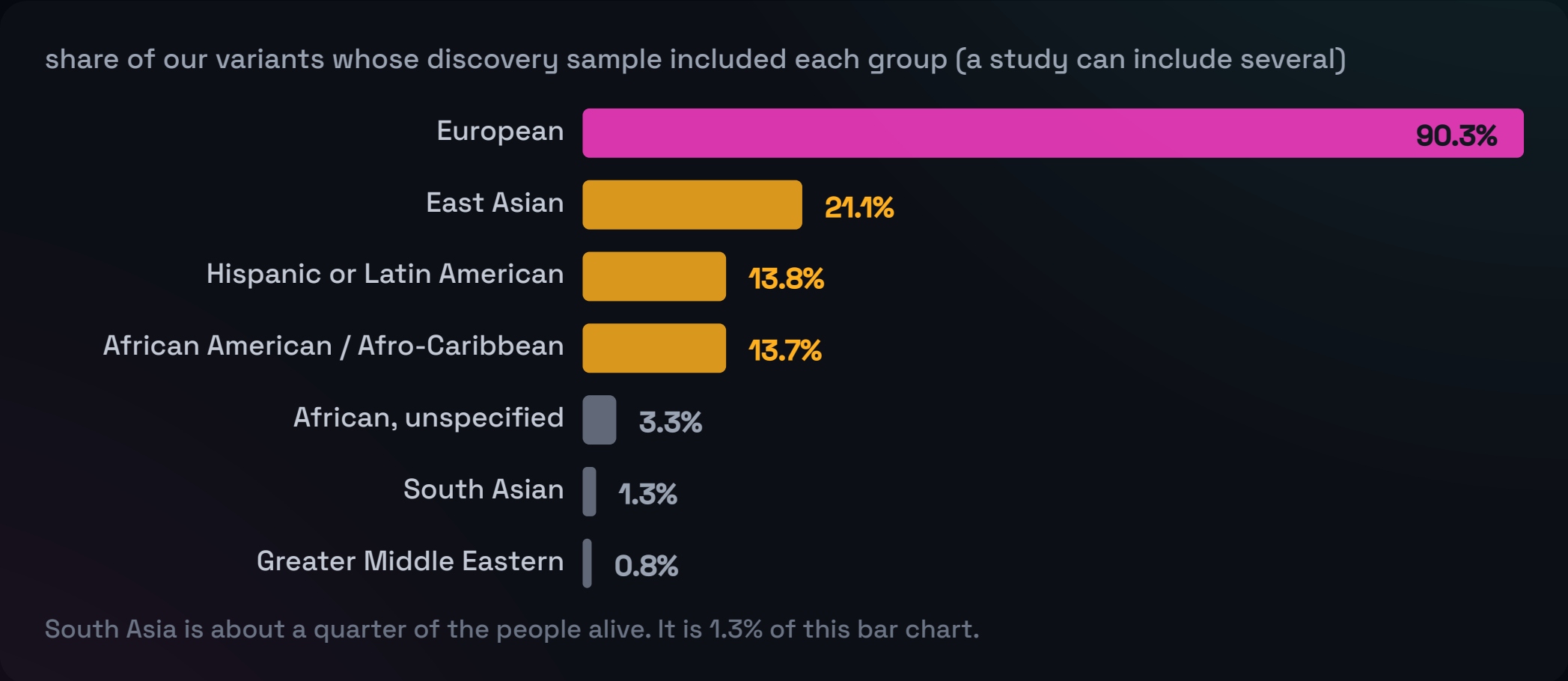
The audit was run on 8 September 2026, against the 608 variants published here then. The

catalogue has grown since and these figures are deliberately not updated with it: the percentages below describe that set of papers, and swapping in a larger number would leave every proportion beside it describing something that was never counted.



Our own catalogue, audited against the GWAS Catalog’s ancestry records for all 187 papers behind it. Not a figure we enjoy publishing, which is most of the reason to publish it.

72.2% of the variants published on this site were discovered in a sample that was entirely European. Another 18.1% came from mixed samples that included Europeans. Fewer than one in ten came from a study with no European ancestry in it at all.



The same audit, cut a second way. Percentages add to more than 100 because one study can recruit several groups — which is what the second bar in the previous figure was.

Cut the other way, it is starker. South Asian ancestry appears in the discovery sample of 8 of our 608 variants — 1.3%. South Asia is about a quarter of the people alive.

We are not the cause of this and we cannot fix it by ourselves: we publish what has been studied, and this is what has been studied. But a site that explains other people’s blind spots while hiding its own is not teaching anything, so the number goes on the page. If a class wants to check it, every source is listed on every variant page and the GWAS Catalog is public.

The argument that is still open

Everyone agrees the imbalance is a problem. What to do about it is genuinely contested, and a class can hold the whole argument.

“Recruit more widely.” The obvious answer, and the right one in the long run. It is also the one with a history: research has been done on communities without their understanding or benefit often enough that suspicion is earned rather than irrational. Any serious answer has to include what participants are told, what they get, and who controls the samples afterwards — not only how many people are enrolled.

“Use better methods instead.” Some argue that statistical methods can borrow strength across populations and partly close the gap without new recruitment. Others answer that a method cannot invent information that was never collected, and that this argument is mostly a way of not doing the expensive thing.

“Then do not use the findings where they do not apply.” Logical, and it has a sting: applied strictly, it would mean withholding genetic risk tools from exactly the populations that have been under-studied, which is a second penalty on top of the first.

There is no consensus to hand a class. Give them the three positions and let them find out that each one has a cost.

Questions

To warm up

1. Why was rs2075650 in this lesson at all, given that nothing interesting happens to it?
2. A reference range for a blood test is built from measurements on healthy people. Give one way that a perfectly correctly built range can still give the wrong answer for a particular patient.
3. What was the largest transfer error you calculated, in percentage points? How did it compare with the guesses on the board?

Core

1. Explain why the Duffy-null calculation used q^2 while the other two used $1 - (1 - q)^2$. What would you need to know about a new variant before choosing between them?
2. Duffy-null became common because it protects against a form of malaria. Explain, in your own words, why a variant can be both strongly beneficial and strongly restricted to one part of the world.
3. 72.2% of this site’s variants were discovered in entirely European samples. Name two separate consequences of that — one for the accuracy of a prediction, and one that is not about accuracy at all.

4. A study reports that a variant is present in 30% of its participants. What is the single most important thing you need to know before using that number anywhere else?

To stretch

1. The labels in these figures — European, East Asian, African — are containers holding a great deal of variation. In the previous lesson on this site, two East Asian populations differed five-fold at one variant. Does that make the labels useless, or useful but insufficient? Defend your answer, and say what you would use instead.
2. Design a rule a hospital could actually follow for deciding when a reference range needs a separate version for a particular group. It has to be specific enough to apply on a Tuesday and it must not require asking every patient about their ancestry.
3. Take the three positions in the argument above. For each, name who bears the cost if it is adopted and who bears the cost if it is not. Which of the three would you defend, and what is the strongest objection to it?
4. This site audited itself and published a number that does not flatter it. Was that the right call? Construct the best argument that it was not — then answer it.

Teacher notes — where each question is going

1. It is the control. Without a case where transfer works, the lesson overshoots into “genetics never generalises”, which is wrong and unhelpful. The method has to be able to return the boring answer.
2. If the healthy people it was built from differ systematically from the patient in something that affects the measurement. Duffy-null is the cleanest example; students may also reach for age, altitude or athletic training, all of which are good.
3. The worst case is 52 points, African versus European at rs9923231. Most classes guess ten or less. Put the two numbers side by side and leave them there.
4. Carrier share is right when one copy is enough to matter; q^2 is right when two are required. Deciding needs the biology — is the effect dominant or recessive — which no amount of allele frequency can tell you. This is the transferable point of the whole lesson.
5. Selection acts where the pressure is. A variant that protects against a parasite becomes common where the parasite is and stays rare where it is not. Good students will connect this to lactase persistence and to why frequency maps often look like maps of something else entirely.
6. Accuracy: predictions and risk scores are less reliable outside the discovery population. Not accuracy: whose questions get asked, which diseases get studied, who benefits from the resulting drugs, and who is expected to trust a

field that has not studied them. Accept any well-argued second consequence.

7. Who the participants were. Everything else — sample size, p-value, effect size — is secondary to whether the number can travel.

8. Useful but insufficient, and the honest answer is that the label is a proxy for the thing that matters, which is genetic ancestry as measured rather than as reported. Strong students may propose using measured ancestry directly, which is what modern studies do, and then run into the fact that this requires genotyping everyone.

9. The hardest practical question here. Watch for rules that quietly require racial classification of patients, and push on them: the Duffy-null case is now often handled by testing the variant itself rather than by asking about background, which is a much better rule and one a class can arrive at.

10. No right answer. The costs are unevenly distributed in all three cases, and noticing that is the outcome.

11. Strongest case against: publishing the weakness invites the finding to be quoted out of context, and could reduce trust in data that is still the best available. The answer we would give is that the number is true, discoverable by anyone with the same public API, and that a reader who finds it themselves after we hid it has learned something much worse about us. Let the class weigh it.

Take it further

- Duffy-null — the full page: the malaria history, the blood counts, and why the trait is not a disorder.
- Warfarin dose variability — why VKORC1 matters clinically, and what genotype does and does not change.
- Lesson: one letter, four subunits — the companion lesson, where two populations a textbook would put in the same box turn out to be five-fold apart.
- Our data page — the live counts, and what actually reads this site.

Use this freely. Print it, copy it, cut it up, put it on your own worksheet, translate it, change the questions. No permission needed and nothing to pay. If you credit it, mygenelog.com is enough.

If you teach with it and something in it does not work, tell us — that is worth more to us than a thank-you.