Blood typing and DNA

Method, formulas, worked example, practice problems, then the lab. No calibration set for this module.

CLAIM CEILING. READ THIS FIRST.

- **Polygraph.** It records breathing and heart rate. A change in either one is a physiological response. It is not a measurement of lying, and nothing in this unit produces a deception call.
- **Hair.** Microscopic similarity can support consistency or support exclusion. It cannot identify one person.
- **Fingerprint.** Pattern class can narrow a field or exclude a person. By itself it cannot name a source, and it never says when or why a finger touched something.
- **Timeline.** A scheduled event is not proof the person was there. Time of death is not established in the early case evidence.

Anna Garcia Forensic Methods Unit. Analysis module.

Every plate, band, gel, measurement, and answer on this card is invented practice data written for this module. None of it comes from the case evidence, and none of it describes any real person.

Vocabulary used on this card.

- **Antigen.** A marker on the outside of a red blood cell. The A antigen, the B antigen, and the D antigen are the three this card uses.
- **Antibody serum.** A liquid that sticks to one specific antigen. Anti-A serum sticks to the A antigen. Anti-B sticks to B. Anti-D sticks to D.
- **Agglutination.** Clumping. When the serum's antibodies grab the matching antigen on many cells at once, the cells stick together and the liquid turns grainy. Clumping means the antigen is present.
- **Rh factor.** The D antigen. Present means Rh positive. Absent means Rh negative.
- **Base pair (bp).** The unit for measuring the length of a DNA fragment.
- **Ladder.** A lane on a gel loaded with DNA fragments of known sizes, run at the same time as the unknowns.
- **Migration distance.** How far a band travelled down the gel from its well, measured in millimetres.
- **Mixture.** A sample containing material from more than one person.
- **Locus.** One specific place on the DNA that a profile looks at. More than one is called loci.

What this analysis can and cannot tell you

- **Claim ceiling.** Blood typing supports two kinds of statement. It can say a person is excluded as the source of a sample. It can say a person is not excluded, and give the fraction of the population that is also not excluded. DNA profiling can narrow much further, but it still describes a sample and a profile, never a person by itself.
- **ABO typing excludes powerfully and includes weakly.** If a stain types B and a person types O, that person did not leave that stain, and that exclusion is strong. If a stain types B and a person types B, you have learned that

about 11 percent of people are still in the running. In a group of 1,000 people, that is about 110 people. Sharing a blood type is common, and common things do not identify anybody.

- **DNA profiling can be far more discriminating, and it depends on sample quality.** A clean, fresh, single-source sample can produce a profile that narrows a field enormously. A small, old, heat-damaged, or bacteria-eaten sample can produce a partial profile that narrows very little, or no usable profile at all. And a **mixture** of two or more people is much harder than a single source, because the extra bands or alleles can be split between the contributors in more than one way, and often more than one combination fits.
- **Neither one tells you when or how a sample got somewhere.** A blood type carries no clock. A DNA profile carries no clock either. Nothing in either result says whether a sample arrived an hour ago or a month ago, or whether it arrived on a hand, on a shoe, on a towel, or on a piece of clothing that changed hands.

Before you start: what you need

For the plate:

This module is on paper. The plates you read are printed inside this booklet, photographed under standard lighting with their controls already run. You need nothing wet.

- This booklet, printed. The plate images are on the lab pages near the back.
- A hand magnifier or a phone camera, so you can look closely at a well that is hard to call.
- A pen. You are recording a reading, and a reading you did not write down did not happen.

If you are also running the real wet lab PLTW Activity 1.1.4 uses a physical typing kit with simulated blood. That is a separate activity with its own materials list from PLTW, and it is not provisioned by this package. Nothing in this booklet needs a plate, a serum, a stirrer or a timer.

For the gel:

- The photographed gel with a visible millimetre scale, or a ruler you can lay on the gel image.
- A calculator with a **log** key and a **10 to the x** key. Check the mode first. Type log 1000. If the screen shows 3, you are fine.
- Graph paper for the standard curve, and a straightedge.
- A sharp pencil.

The formulas, explained

1. ABO agglutination logic

The rule. Clumping in a well means that antigen is on the cells. No clumping means that antigen is absent. Three wells give three yes-or-no answers, and three yes-or-no answers give eight possible outcomes. Every one of them is a real blood type, so read the table, do not reason it out from memory.

Anti-A well	Anti-B well	Anti-D well	Blood type
clumps	no clumps	clumps	A positive
clumps	no clumps	no clumps	A negative
no clumps	clumps	clumps	B positive
no clumps	clumps	no clumps	B negative
clumps	clumps	clumps	AB positive
clumps	clumps	no clumps	AB negative
no clumps	no clumps	clumps	O positive
no clumps	no clumps	no clumps	O negative

Read it as three columns in, one answer out. The anti-A and anti-B wells together give the letter. The anti-D well alone gives the positive or negative.

What "clumps" looks like. Grains, specks, or small clots suspended in a liquid that has gone clearer around them. Tilt the plate and the grains move as lumps. **What "no clumps" looks like.** An even, smooth, cloudy red liquid all the way across. Tilt it and it flows as one thing.

Assumptions built in. The sera are fresh and working. The sample has intact red cells. The sample came from one person.

When it breaks. It breaks on old or dried sera, which can fail to clump a sample that should clump, and read as a false negative. It breaks on a sample so degraded that the cells have already broken open, because there is nothing left to clump. It breaks on a mixture, and this failure is the dangerous one. A mixture of a type A person's blood and a type B person's blood will clump in the anti-A well and clump in the anti-B well, which is exactly the reading a single type AB person gives. The plate cannot tell those two situations apart.

How the plate you are reading was produced. You are not running this. Read it anyway, because you cannot judge a result whose procedure you do not know. The technician put the plate on white paper, added one drop of sample to each of the three wells, added one drop of the matching serum to each, and stirred each well with its own clean stirrer, throwing each one away, because a stirrer carried between two wells contaminates both. They rocked the plate gently for two minutes, then read all three wells against white paper and again with the light behind them. They read two control samples the same way. The known A must clump in anti-A and not in anti-B. The known O must clump in neither. If either control does the wrong thing the plate is void, and no reading from it can be used.

At your desk. Find the printed plate for your sample on the lab pages. Read the two control wells FIRST and write down what each one did. If the controls did not behave, stop and write "plate void" and nothing else. Only when both controls are correct, read the three sample wells and write each one down as "clumps" or "no clumps". Then read the blood type straight off the table.

You are done with this step when: both control results are written down and both are correct, all three wells are written down as "clumps" or "no clumps," and one blood type is read straight off the table.

2. Population frequency

The numbers. These are approximate teaching values for the United States as a whole. They are rounded, and they vary a great deal between ancestry groups, so a number from this table is a rough guide for a classroom exercise and is not a figure to quote as a fact about any particular group of people.

Type	Approximate United States frequency
O positive	about 37 percent
A positive	about 36 percent
B positive	about 9 percent
O negative	about 6 percent
A negative	about 6 percent
AB positive	about 3 percent
B negative	about 2 percent
AB negative	about 1 percent

Adding the positive and negative versions together gives the letter frequencies. O is about 43 percent. A is about 42 percent. B is about 11 percent. AB is about 4 percent.

The arithmetic. To find how many people in a group of size N would be expected to share a type by chance, multiply.

> expected number = frequency as a decimal x N

What it means for narrowing a field. Every person in your field whose type differs from the stain's type is **excluded**, and that is real, useful work. Every person whose type matches is **not excluded**, and that is a much weaker statement. Report both halves with the denominator attached. "Four of the five are excluded, one is not excluded" is honest. "It matches suspect three" is not.

When it breaks. It breaks if you treat the number as though it applies to one person. A frequency of 11 percent does not mean an 11 percent chance that a particular person left the stain. It means about 11 out of every 100 people would also have failed to be excluded. It also breaks if you quote a nationwide number as though it described a specific community, because these frequencies shift substantially between ancestry groups.

At your desk. Write the stain's type at the top of your sheet. Write every person in your field underneath with their type. Cross out every one that differs. Count how many are crossed out and how many remain. Then multiply the type's frequency by 1,000 and write that number down as well, so the size of the group you have not excluded is on the page next to the small field you were handed.

You are done with this step when: the number excluded, the number not excluded, the total field size, and the expected count in a group of 1,000 are all written on your sheet.

3. Gel electrophoresis sizing

Why DNA moves at all. The backbone of a DNA molecule carries phosphate groups, and phosphate groups are negatively charged. Opposite charges attract, so in an electric field every DNA fragment is pulled toward the **positive** electrode. Load the wells at the negative end and the DNA runs toward the positive end. Get this backwards when you set up and the DNA runs out of the gel into the buffer and is gone.

Why size matters. The gel is a mesh. Small fragments slip through the mesh easily and travel far. Large fragments snag and travel a short distance. So

smaller fragments move further, and distance travelled is a clue to size.

The formula.

> distance = m x log10(size) + b

What every symbol means.

- **distance** is the migration distance in millimetres, measured from the well to the middle of the band.
- **size** is the fragment length in base pairs.
- **log10(size)** is the power you raise 10 to in order to get that size. log10 of 1000 is 3, because 10 x 10 x 10 = 1000. log10 of 100 is 2. For sizes that are not neat powers of ten, use the calculator's log key.
- **m** is the slope of the line, in millimetres per log unit. It comes out **negative**, because bigger fragments travel less far.
- **b** is the intercept, in millimetres. It is where the line would sit at log10(size) = 0.

Why a log and not the size itself. Plot distance against raw size and you get a curve that bends hard, and you cannot read anything off a bent line with a ruler. Plot distance against log10 of size and the points fall very close to a straight line over the useful range. A straight line is something you can draw, extend, and read.

How you build the line. You do not know m and b in advance. You get them from the **ladder**, a lane loaded with fragments of known size run on the same gel at the same time. Measure each ladder band's distance, take log10 of each known size, plot distance against log10 of size, and draw the best straight line through the points. Then read each unknown band off that line.

Assumptions built in. The unknown lane and the ladder lane ran on the same gel, at the same voltage, for the same length of time. The gel is even across its width. The unknown's size falls inside the ladder's range.

When it breaks. It breaks if there is no ladder, because then there is no line and a distance in millimetres means nothing at all. It breaks if the unknown band is bigger than the largest ladder band or smaller than the smallest, because reading past the end of your line is guessing. It breaks completely if a band ran

off the end of the gel, because a band that left the gel has no measurable distance, so it has no size. "Smaller than the last band still on the gel" is the only honest thing you can say about it.

At your desk. Lay the ruler along the ladder lane with zero at the well. Read each band's distance to the middle of the band, in millimetres. Take log10 of each known size. Plot the points on graph paper with log10 of size on the bottom and distance up the side, and draw one straight line through them with the straightedge. Then measure each unknown band's distance the same way, find that distance on the side of your graph, run across to the line, drop down, read log10 of size, and press the 10 to the x key to turn it back into base pairs.

You are done with this step when: every ladder band has a distance and a log10 value, a straight line is drawn through the plotted points, and every unknown band has a distance, a log10 value read off the line, and a size in base pairs.

Worked example, done all the way through

Luis is working a practice set called PRACTICE-BT. It is a made-up plate and a made-up gel written for this module. None of it is case evidence.

Part one, the plate

Luis puts plate PRACTICE-BT-1 on white paper, loads the three wells, uses a fresh stirrer in each, and rocks it for two minutes.

Controls first. The known type A control clumps hard in anti-A and stays smooth in anti-B. The known type O control stays smooth in both. Both controls did what they should, so the sera are working and the plate can be read.

- **Anti-A well.** Smooth, even, cloudy red. Tilting it, the whole thing flows as one liquid. **No clumps.**
- **Anti-B well.** Clear grains suspended in a fluid that has gone thinner around them. Tilting it, the grains move as lumps. **Clumps.**
- **Anti-D well.** This one stops Luis. Held flat against the white paper it looks smooth. Held up to the light it looks very slightly grainy near one edge. It is not the clean grain of the anti-B well and it is not the clean smooth of the anti-A well.

Luis's first instinct is to call it a clump, because two out of three wells are clear and it feels wrong to leave one blank. Under the table that reading gives B positive, which is a tidy answer.

Instead Luis runs the anti-D well again with a fresh drop of sample, a fresh drop of serum, a new stirrer, and the two controls beside it. The known Rh positive control clumps strongly, so the anti-D serum is fine. The repeat of the sample comes out the same as before, faintly and unevenly grainy near one edge.

So the serum works, the technique works, and the sample still will not give a clean answer. Luis writes it down that way.

> PRACTICE-BT-1. Anti-A no clumps. Anti-B clumps. Anti-D ambiguous on two independent readings, with the Rh positive control clumping strongly on both.
> ABO type is **B**. Rh factor **not determined**.

"Not determined" is a real result and it is the correct one here. Guessing "positive" to fill the box would have put a number in the record that the plate never produced.

Part two, how much does that narrow a field of five?

The five practice individuals in this exercise type as follows. These are invented for the exercise.

Practice individual	Type
PRACTICE-S1	O positive

Practice individual	Type
PRACTICE-S2	A positive
PRACTICE-S3	B positive
PRACTICE-S4	O positive
PRACTICE-S5	A negative

The stain is type B. S1, S2, S4, and S5 are types O, A, O, and A. **Four of the five are excluded.** S3 is type B, so S3 is **not excluded**.

That leaves exactly one person in this small field, and it is tempting to write that the stain came from S3. It does not follow. Type B is about 11 percent of the population. In a group of 1,000 people, about 110 would also fail to be excluded. The field of five was handed to Luis by the exercise. The blood type did not choose S3 out of the world. It removed four people from a list of five that somebody else wrote.

The undetermined Rh result costs something real too, and Luis writes that down. S3 is B positive. If the anti-D well had read cleanly negative, S3 would have been excluded as well and the field would have been empty. An honest "not determined" gave up exclusionary power that a guess would have faked.

Part three, the gel

Gel PRACTICE-GEL-1 has a ladder lane and two unknown bands, U1 and U2. Luis lays the ruler with zero at the wells and reads to the middle of each band.

Ladder band	Known size (bp)	log10(size)	Distance (mm)
L1	10000	4.0000	12
L2	3000	3.4771	22
L3	1000	3.0000	31
L4	500	2.6990	37
L5	100	2.0000	50

Luis plots distance up the side against log10 of size along the bottom, and the five points fall close to a straight line sloping downward.

Finding the line. Luis uses the two end points, L1 and L5, because they are furthest apart and a line through two far-apart points is the least wobbly.

> m = (50 minus 12) divided by (2.0000 minus 4.0000)
> m = 38 divided by negative 2.0000
> m = **negative 19.0 mm per log unit**

The slope is negative, which is what it should be, because distance goes down as size goes up.

> b = 12 minus (negative 19.0 x 4.0000)
> b = 12 plus 76
> b = **88.0 mm**

So the standard curve is **distance = negative 19.0 x log10(size) + 88.0**.

Check the line against the three middle bands. This is the part students skip and it is the part that catches a bad line.

- L2, 3000 bp: negative 19.0 x 3.4771 is negative 66.1, plus 88.0 is 21.9 mm. Measured 22 mm. Off by 0.1 mm.
- L3, 1000 bp: negative 19.0 x 3.0000 is negative 57.0, plus 88.0 is 31.0 mm. Measured 31 mm. Off by 0.
- L4, 500 bp: negative 19.0 x 2.6990 is negative 51.3, plus 88.0 is 36.7 mm. Measured 37 mm. Off by 0.3 mm.

Three known bands land within 0.3 mm of the line. The line is good.

Reading the unknowns. Rearrange the formula to solve for log10 of size.

> log10(size) = (88.0 minus distance) divided by 19.0

Unknown U1, distance 27 mm.

> log10(size) = (88.0 minus 27) divided by 19.0
> log10(size) = 61 divided by 19.0 = 3.2105
> size = 10 to the 3.2105 = **about 1,624 bp**

Unknown U2, distance 44 mm.

> log10(size) = (88.0 minus 44) divided by 19.0
> log10(size) = 44 divided by 19.0 = 2.3158
> size = 10 to the 2.3158 = **about 207 bp**

Both distances fall between ladder bands, so both are genuine interpolations and not guesses off the end of the line.

Part four, the second moment Luis gets stuck

Luis writes "U1 = 1,624 bp" and stops, because 1,624 looks like a measurement of a thing, and it is not. It came from a pencil line on graph paper and a ruler read by eye.

So Luis tests it. Suppose the distance reading was off by 1 mm in either direction, which is about as well as anyone reads a band edge by hand. One millimetre divided by 19.0 is 0.0526 log units. Turning that back into a size factor, 10 to the 0.0526 is 1.129. So a 1 mm slip multiplies or divides the answer by about 1.13.

- U1: 1,624 divided by 1.129 is 1,438. 1,624 times 1.129 is 1,834. So U1 is somewhere around **1,440 to 1,830 bp**.
- U2: 207 divided by 1.129 is 183. 207 times 1.129 is 234. So U2 is somewhere around **183 to 234 bp**.

Writing "1,624 bp" claims the band is known to the nearest base pair. Nothing on that gel supports that. Luis reports ranges.

> PRACTICE-GEL-1, read against a five-band ladder on the same gel. Standard curve distance = negative 19.0 x log10(size) + 88.0, checked against three ladder bands and accurate to within 0.3 mm. Band U1 at 27 mm sizes to roughly 1,440 to 1,830 base pairs. Band U2 at 44 mm sizes to roughly 183 to 234 base pairs. Both bands fall inside the ladder's range. These are fragment sizes. They are not a profile and they do not name anyone.

Environmental factors that change the answer

Factor	What it does to the result	What you must record
Sample age	Older samples give weaker agglutination and shorter, fewer DNA fragments. Very old samples may give nothing readable at all.	Date and time collected if known, and how long the sample sat before testing.
Degradation	DNA breaks into shorter and shorter pieces over time. On a gel that shows up as a smear instead of clean bands, and larger bands vanish first.	Write "smeared" or "bands clean," and note whether the large bands are missing.
Heat	Heat speeds up every kind of breakdown. A sample left in a hot car or in sun degrades far faster than the same sample kept cool.	Where the sample was stored, and the rough temperature.
Humidity and moisture	Damp keeps a sample soft and lets bacteria grow. Dry storage slows both. A rewetted dry stain is not the same as a fresh one.	Damp or dry, and whether the stain was ever wet after it dried.
Bacterial contamination	Bacteria eat the sample and add their own DNA. That can produce extra bands that belong to nothing you care about, and it destroys red cells so the plate stops clumping.	Any smell, discolouration, mould, or slime, and where the sample was found.
Mixed sample from more than one person	On a plate, a type A person plus a type B person reads exactly like one type AB person. On a gel, you get more bands than one source should give, and more than one combination of contributors can fit.	Anything suggesting more than one source. Never assume one source just because you were handed one swab.
Sample volume	Too little sample gives faint agglutination that is easy to misread, and a faint gel band that is easy to miss. Too much can look like clumping when it is only settled cells.	The size of the stain or the volume of the sample, in your own units.
Porous surface (cloth, carpet, unsealed wood, paper)	The sample soaks in and mixes with whatever the material carries. Recovery is lower and contamination is higher than from glass or tile.	The surface material, and whether the stain was on it or in it.
Gel running time	Run longer and every band moves further. The same fragment sits at a different distance on a gel run for 30 minutes than on one run for 60.	Running time, and always run the ladder on the same gel.

Factor	What it does to the result	What you must record
Gel voltage	Higher voltage moves bands faster and can blur them. Lower voltage is slower and sharper. Either way the distances shift.	Voltage, and always run the ladder on the same gel.
A band that ran off the end of the gel	It has no measurable distance, so it has no calculable size. The formula needs a distance, and there is not one.	Write "ran off the gel, size not determinable, smaller than the last visible band." Rerun for a shorter time if you can.

The last three rows are the reason a ladder is not optional. Running time and voltage change every distance on the gel. Running the ladder on the same gel, at the same time, at the same voltage, cancels all of that, because the ladder shifts by exactly as much as the unknowns do.

The logic a real examiner uses

Exclusion is strong and inclusion is weak. These are not two sides of the same result. If the stain types B and the person types O, that person is out, and that conclusion is about as solid as this method gets. If the stain types B and the person types B, all you have learned is that the person survived a filter that about 11 percent of everybody survives. Write exclusions with confidence. Write inclusions with the fraction attached.

Report how much a result narrows the field, with the denominator. A number without a denominator is not a result. "Four of five excluded, one not excluded, and about 11 in every 100 people would also not be excluded" tells the reader exactly what happened. "It matches" tells the reader nothing and suggests something the test never delivered.

Never say a blood type identifies a person. The word "match" does the damage. It carries a suggestion of one-to-one, and a blood type is one-to-millions. Retire it. The sentences that stay inside the ceiling are "excluded as the source of this sample" and "not excluded as the source of this sample."

Say what a DNA profile covers and what it does not. A profile reads a fixed set of specific locations on the DNA, called loci. It does not read the whole genome, and it is not a photograph of a person. So a report has to say how many loci were read, how many gave a clean result, and how many failed. A profile that covers every locus tested narrows a field enormously. A **partial profile**, where several loci failed because the sample was small or degraded, narrows far less, and how much less depends on which loci survived. Two people who cannot be told apart on four surviving loci could be told apart easily on all of them. Report the count. A profile described without saying how complete it is has hidden the part that decides what it is worth.

Treat every sample as a possible mixture until you have a reason not to. One swab does not mean one person. Check the plate reading against the possibility of two contributors. Count the bands on the gel against what a single source should give. If more than one combination of people fits your data, say so, and say how many fit.

Run your controls and report them. A known positive and a known negative on every plate, and a ladder on every gel. A result with no control behind it cannot be told apart from an equipment failure. If a control fails, the run is void, and that is what you write.

Practice problems

All data below is invented practice data. Work each one on paper before you look at the answers. Problems 3 and 4 use the standard curve from the worked example,

distance = negative 19.0 x log10(size) + 88.0, which rearranges to

log10(size) = (88.0 minus distance) divided by 19.0.

Problem 1. A plate reads as follows. Anti-A well, clumps. Anti-B well, no clumps. Anti-D well, no clumps. Both controls behaved correctly. What is the blood type?

Problem 2. A second plate reads as follows. Anti-A well, clumps. Anti-B well, clumps. Anti-D well, clumps. Both controls behaved correctly. What is the blood type? Then name one other situation that would produce exactly this reading, and say how you would tell the two apart from the plate alone.

Problem 3. An unknown band on PRACTICE-GEL-1 sits 34 mm from the well. Calculate its size in base pairs. Then give the range you would actually report if your distance reading could be off by 1 mm.

Problem 4. A second unknown band on the same gel sits 19 mm from the well. Calculate its size in base pairs, and give the reported range for a 1 mm reading error. Then say whether this band is inside the ladder's range.

Problem 5. A stain types O positive. The field handed to you has 8 people in it. Using the frequency table, how many of those 8 would you expect to type O positive purely by chance? What does that tell you about the strength of an "included" result here? Support your answer with a number for a group of 1,000 people.

Problem 6. A dried stain is scraped off a dashboard after a week in a car parked in the sun. It is run on a gel next to a fresh ladder. The ladder lane gives five sharp, well-separated bands at the expected positions. The sample lane gives a continuous grey smear running from the well down to about 45 mm, with no distinct band anywhere in it. What do you report, and why does the ladder lane matter to your answer?

Answers to the practice problems

Answer 1.

Anti-A clumps, so the A antigen is present. Anti-B does not clump, so the B antigen is absent. That gives the letter **A**. Anti-D does not clump, so the D antigen is absent, which is **negative**.

The type is **A negative**. From the frequency table, that is about 6 percent of the population.

Answer 2.

All three wells clump. A antigen present, B antigen present, D antigen present. Read straight off the table, the type is **AB positive**, about 3 percent of the population.

The other situation that gives exactly this reading is a **mixture**. Blood from a type A person combined with blood from a type B person will clump in the anti-A well, because the A cells are there, and clump in the anti-B well, because the B cells are there. The plate has no way to know whether the A antigen and the B antigen arrived on the same cells or on different cells from different people.

You cannot tell the two apart from the plate alone. That is the whole answer, and writing anything else is the mistake. What you do instead is report it as "consistent with AB, and also consistent with a mixture of a type A source and a type B source," and note that further testing would be needed to separate them. AB is the rarest type on the table at about 4 percent for both Rh versions combined, so an AB reading is exactly the reading where you should slow down and ask about a mixture, not speed up because the answer looks rare and therefore useful.

Answer 3.

> log10(size) = (88.0 minus 34) divided by 19.0
> log10(size) = 54 divided by 19.0 = 2.8421
> size = 10 to the 2.8421 = **about 695 bp**

Now the reported range. A 1 mm error is 1 divided by 19.0 = 0.0526 log units, and 10 to the 0.0526 is a factor of 1.129.

- 695 divided by 1.129 = 616
- 695 times 1.129 = 785

Report the band as **roughly 616 to 785 base pairs**. It sits between the 1000 bp ladder band at 31 mm and the 500 bp band at 37 mm, so it is inside the ladder's range.

Answer 4.

> log10(size) = (88.0 minus 19) divided by 19.0
> log10(size) = 69 divided by 19.0 = 3.6316
> size = 10 to the 3.6316 = **about 4,282 bp**

Range using the same factor of 1.129.

- 4,282 divided by 1.129 = 3,792
- 4,282 times 1.129 = 4,834

Report the band as **roughly 3,792 to 4,834 base pairs**. When you round an uncertainty range, always round outward. Rounding inward makes your result look more precise than your ruler was.

Is it inside the ladder's range? Yes. The ladder runs from 12 mm (10000 bp) down to 50 mm (100 bp), and 19 mm sits between the 12 mm band and the 22 mm band. So this is an interpolation between two known points, which is allowed. If the band had sat at 8 mm, above the largest ladder band, you would be reading past the end of your line and the honest answer would be "larger than 10000 bp, size not determinable."

Notice also that the band's range is about 1,000 bp wide, which is much wider in raw base pairs than the range in answer 3. That is what a log scale does. The same 1 mm of ruler error is worth far more base pairs up at the large-fragment end of the gel, where the bands are crowded together.

Answer 5.

O positive is about 37 percent. As a decimal that is 0.37.

> expected number = 0.37 x 8 = 2.96, so **about 3 of the 8**

So before you test anybody, you should expect roughly 3 people in a random group of 8 to type O positive. The test will exclude the other 5 or so, and that exclusion is genuinely useful. What it leaves behind is a group, not a person.

For a group of 1,000 people, 0.37 x 1,000 = **370 people** would also fail to be excluded.

An "included" result here narrows almost nothing. O positive is the single most common type on the table. Being O positive puts a person in the largest group there is. The honest report is that about 5 of the 8 are excluded, about 3 are not excluded, and roughly 370 people in every 1,000 would also not be excluded.

Answer 6.

Report no result. Write it as "sample lane shows a continuous smear from the well to approximately 45 mm with no resolvable bands. No fragment sizes can be determined from this lane."

Do not pick a darker spot in the smear and read a size off it. A smear is what degraded DNA looks like, and it means the sample has broken into fragments of every length at once instead of a few defined lengths. Heat, a week of time, and sun are exactly the conditions that produce it. Choosing a point inside a continuous smear and calling it a band is choosing your own answer, because you could have chosen any point and got any size.

Why the ladder lane matters. The ladder is your positive control on the whole run. It gave five sharp, well-separated bands at the expected positions, which proves the gel poured correctly, the voltage was right, the running time was right, the buffer was fine, and the electrodes were the correct way round. Every part of the equipment worked. So the smear cannot be blamed on the gel, and there is nothing to fix by rerunning it. The problem is in the sample, and no rerun of a degraded sample makes it less degraded. If the ladder lane had also been smeared, you would have a completely different conclusion, which is that the run failed and the sample has not been tested yet.

The three mistakes that cost people most

Treating a shared blood type as an identification. This is the most common error and it hides inside ordinary words. A student writes "the stain matches suspect three," and the sentence sounds careful, and it is not. The stain typed B. Suspect three typed B. About 11 percent of people type B, which is roughly 110 people in every 1,000 and roughly 110,000 in every million. The test did not select suspect three. It failed to remove suspect three from a list somebody else wrote. Those are different results and only one of them is what happened. The fix is to stop writing the word "match" entirely. Write "excluded as the source of this sample" or "not excluded as the source of this sample," and put the population fraction beside every inclusion, every time. If you find yourself wanting to leave the fraction off because it makes the result look weak, that is the moment to put it in, because the result **is** that weak and the reader needs to know.

Reading a gel without a ladder. A distance in millimetres is not a size. It is a distance. The same 1,000 bp fragment lands at 31 mm on the gel in this module and somewhere else entirely on a gel run for twice as long, or at a different voltage, or in a differently poured gel. Without a ladder run on that same gel at that same time, there is no slope, no intercept, no line, and nothing to read an unknown against. Students still try, usually by comparing the unknown band to a picture of a gel from a different day. That comparison has no arithmetic behind it. And a ladder does a second job that is just as valuable. It is your positive control, as answer 6 shows. When the ladder comes out clean, you know the equipment worked and any problem lives in the sample. When the ladder comes out wrong, the whole run is void, and knowing that saves you from reporting a confident number produced by a broken gel.

Ignoring the possibility of a mixture. Every method on this card was designed around one assumption, which is that the sample came from one person. Nothing on the plate and nothing on the gel checks that assumption for you. Two people can produce a plate reading that is indistinguishable from one person's, and answer 2 shows the exact case, where type A blood plus type B blood reads as AB. On a gel, a mixture puts extra bands in the lane, and there is usually more than one way to split those bands between two contributors, so more than one pair of people fits the same data. The trap is that a mixture almost never announces itself. It looks like a clean, ordinary, single-source result, and it is the results that look cleanest that get reported with the most confidence. So make the check a step rather than an instinct. Before you write any conclusion, ask out loud whether two sources could have produced this exact reading. If the answer is yes, that sentence goes in the report next to your result, and it goes in whether or not you think a mixture is likely.

The lab. Blood typing and a gel

Practice evidence for an invented case. Read the wells first, before you look at any label.

Agglutination plate. Read the wells, do not read the labels.

Clumping means the antibody found its antigen. A smooth, even well means it did not.

Sample	Anti-A	Anti-B	Anti-D	Type
Scene stain S1				
Scene stain S2				
Person 1				
Person 2				

Sample	Anti-A	Anti-B	Anti-D	Type
Person 3				
Person 4				
Person 5				

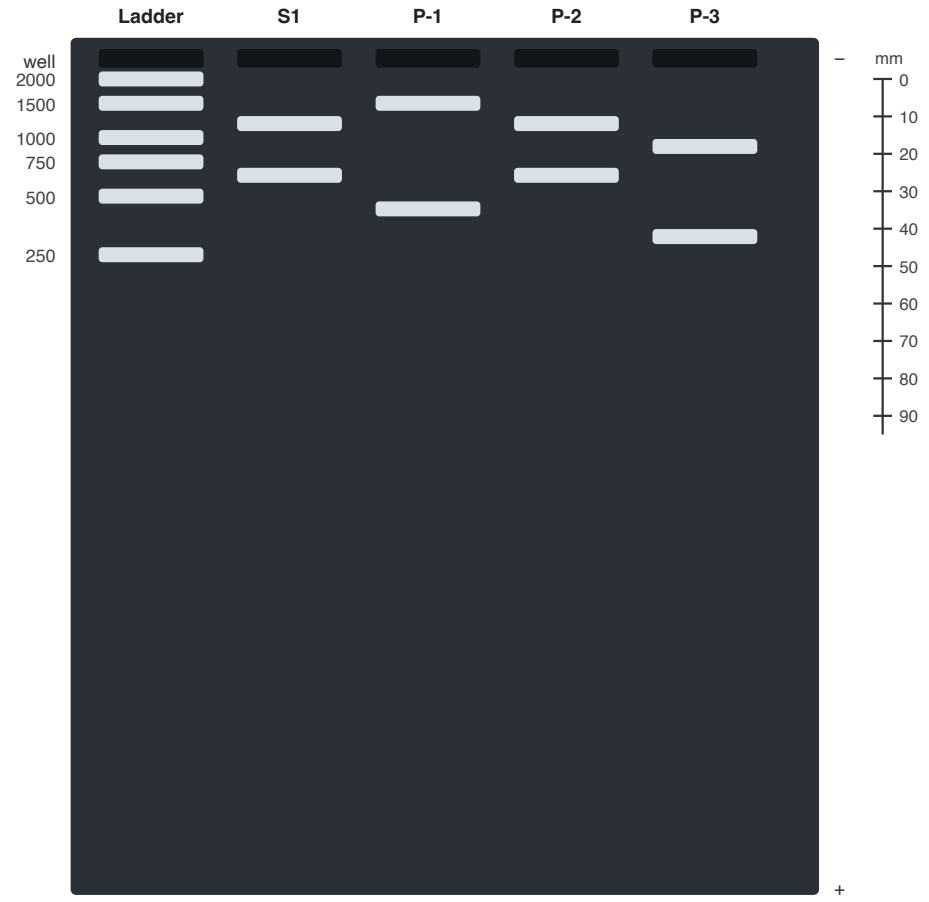
How much did that narrow the field?

Scene stain S1 has a type. How many of the five people share it? Write the fraction, then write what that fraction does and does not support.

Sizing the gel

Build a standard curve from the ladder, then read the unknowns off it.

Electrophoresis gel. Ladder sizes are in base pairs.



1. Measure how far each ladder band travelled, in millimetres, using the scale beside the gel.
2. Take log base 10 of each ladder size in base pairs.
3. Plot distance against log10(size). Draw the best straight line.
4. Measure each unknown band, find it on your line, and read off log10(size). Undo the log to get base pairs.

Ladder band	Distance (mm)	log10(size)
Ladder 2000		
Ladder 1500		
Ladder 1000		
Ladder 750		
Ladder 500		

Ladder band	Distance (mm)	log10(size)
Ladder 250		

Unknown	Distance (mm)	log10(size) from your line	Size (bp)
S1 upper band			
S1 lower band			
P-1 upper			
P-1 lower			
P-2 upper			
P-2 lower			
P-3 upper			
P-3 lower			

Every lane has two bands. Record both. A lane only matches the scene stain if BOTH bands agree, and a single band in common is exactly the kind of partial agreement that gets over-read.

> WHY THE LOG
>
> Distance is a straight line against **log** of size, not against size. Plot distance against raw size and you get a curve you cannot read. That is why every gel needs a ladder run on the same gel, at the same voltage, for the same time.

Write the conclusion

Which person's lane matches the scene stain? Write one sentence on what that supports and one on what it does not. Neither sentence may use the words match, identified, or proves.