Antibody identification through software.

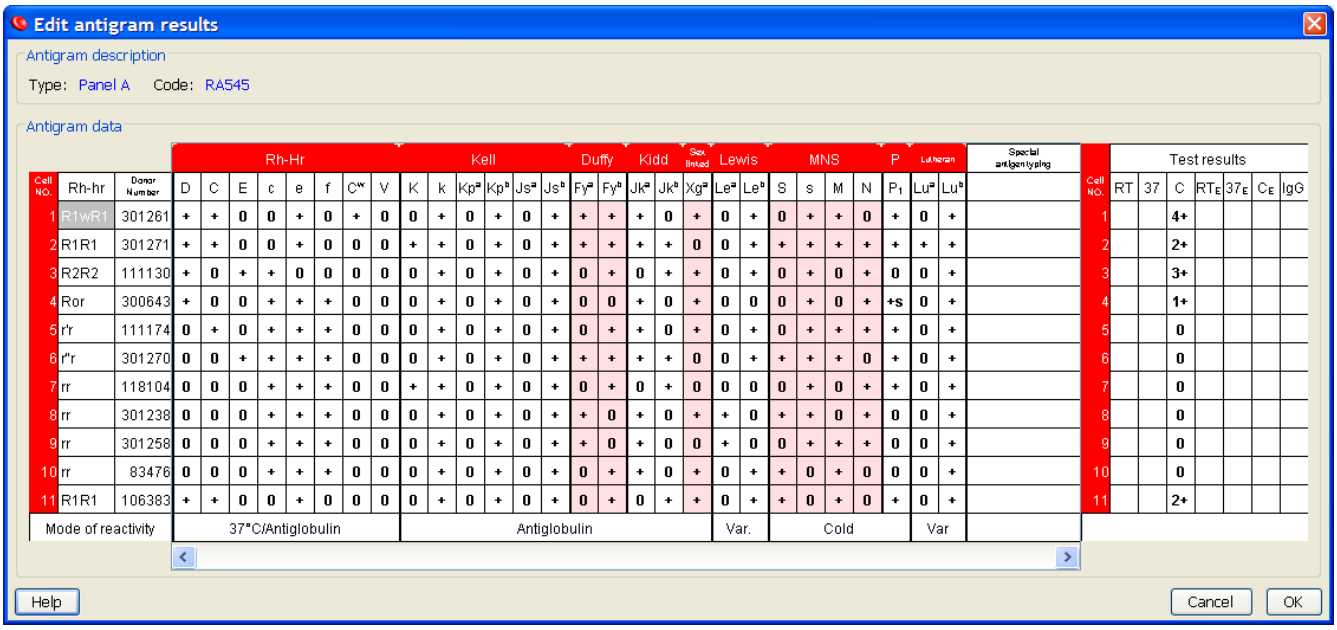
By Carlo Magnaghi – 27 March 2009

Introduction

The principles on which an antibody identification software is based are not substantially different from the way a human would proceed for the same task, therefore I will restate those principles and show how they are executed from the software viewpoint.

Working by hand

The first step for identifying antibodies with a diagnostic panel, after entering scores obtained manually or importing results from an automatic machine such as the Autovue Innova, is excluding all antibody specificities whose target antigen is expressed on a non reacting RBCs sample.



With the above results we can exclude the presence of the following specificities:

- anti-C, anti-c, anti-e, anti-f, anti-k, anti-Kp^b, anti-Js^b, anti-Fy^b, anti-Jk^a, anti-Xg^a, anti-Le^b, anti-s, anti-M, anti-N, anti-P_1 and anti-Lu^b, because RBCs sample 5 is not reacting.
- Anti-E, anti-Fy^a, anti-Jk^b, anti-S, because RBCs sample 6 is not reacting.
- Anti-K, because RBCs sample 7 is not reacting.
- Anti-Le^a, because RBCs sample 8 is not reacting.

We are thus left with the following antibody specificities: anti-D, anti-C^w, anti-V, anti-Kp^a, anti-Js^a, anti-Lu^a.

A not rare misconception is that we can also exclude the antibody specificities whose antigen is not expressed on a reacting RBCs sample (like anti-C^w on RBC 2 above). This is only true if we assume that there is ONLY one antibody specificity in patient's serum, which is indeed the next most common case (after no antibody specificity reacting with RBCs on panel, which is fortunately the rule), but not the only case.

If we used this rule we could exclude anti-C^w, anti-V, anti-Kp^a, anti-Js^a and anti-Lu^a, leaving anti-D as the only possibility. This would bring us very close to the true answer in this case, but could be totally misleading in other cases.

What we can say is that anti-D alone is sufficient to justify the reactivities (as far as we neglect the scores) and that the other antibody specificities are possible.

At this point we should enter in an other line of reasoning, based on the likelihood and clinical significance of the different non-excluded antibodies.
If you feel uneasy about excluding antibody specificities not based on hard evidence, you should consider that the antigenic expressions listed on diagnostic panels are by no means a complete list of all the antigens for which it has been observed that an antibody can be developed, such a list including several hundreds of antigens.

Using this somewhat relaxed sort of reasoning we could conclude that:

- anti-D alone is enough to justify all the reacting RBCs.
- Anti-C^w is not needed (more on this later).
- anti-V cannot be detected using this panel; so if we consider anti-V as a potential problem we need to add V+ RBCs.
- Anti-Js^a cannot be detected using this panel; so if we consider anti-Js^a as a potential problem we need to add Js(a+) RBCs.
- Anti-Kp^a cannot be detected using this panel; so if we consider anti-Kp^a as a potential problem we need to add Kp(a+) RBCs.
- Anti-Lu^a cannot be detected when anti-D is present (anti-D reactions 'cover' anti-Lu^a. reactions); hence if detecting Lu^a is important we should add D- Lu(a+) RBCs.

Thus a first identification could be anti-D.

If we now examine scores more closely we can observe that sample 1 is reacting more strongly that sample 3; this is strange, because R2R2 RBCs have the strongest expression for the D antigen and R1wR1 should react more weakly; this kind of dose (haplotype) effect is not always noticeable, but here it is reversed, which should ring a bell. This should be enough evidence to conclude that anti-C^w specificity is likely present in the antibody mix.

Can something also be said for anti-Lu^a? Actually we see that the only Lu(a+) sample, namely sample 2, is R1R1 and is reacting exactly as the other R1R1 sample (number 11). So not only there is no evidence of an anti-Lu^a antibody, but, if present it must react weakly enough that it does not change the score on sample 2. Further anti-Lu^a antibodies have never been implicated in transfusion reactions and are considered not clinically significant.

With this information we could conclude that the reactions are most likely due to a mix of anti-D and anti C^w.

Having found an additional antibody specificity just by examining scores, we could suspect that something more is hidden: after all the four D+ C^w- RBC samples range in score from 1+ to 3+. We need to look at the Rh haplotypes for an explanation for this: we know that the expression of the D antigen on D+ RBCs follows the pattern R2R2>R1R1>R1wR1>R0r. Hence what we are observing is just a pattern of reactions that matches the D antigen expressions. Note that this haplotype effect is not always so strongly marked by different anti-D antibodies.

In this simple example we have seen a lot of what goes into the task of antibody identification; in particular we have seen that identifying antibodies is not a 'MasterMind' like activity involving only combinatorial analysis: proficiency in the immuno-hematological characteristics and clinical significance of the antibodies is paramount.
Quick and dirty approaches, such as using the Fisher formula, are too crude to tackle the intricacies involved: Anti-D, Anti-D+anti-C^w and anti-D+anti-Lu^a would give the same Fisher formula results,

yet we have seen that there is strong evidence to favor one of these combinations over the others.

The computer based approach

The first step taken by a computer to analyze reactions on a given panel or set of panels is to build a table of the antibodies that could possibly be present with the scores they could yield.
In the example above we could create a table like the following.

	Reaction with RBC samples where antigen is		
Antibody	***Weak***	***Average***	***Strong***
Anti-D	1	2	3
Anti-C	0	+/-	1
Anti-E	0		+/-
Anti-c	0	0	+/-
Anti-C^w		4+	
Anti-P1			+/-
Anti-Lu^a	2+		

Note that some of the antibody specificities that we considered possible in human analysis (anti-V and anti-Js^a) are not listed: the fact is that no analysis is possible for these antibodies because they are not present on the panel so it would be pointless to include them in the list. What a software can do is to check if the antibodies that are not present on panel are frequent enough or clinically significant and post a warning to the user who should evaluate if further tests are required.
The rule of the game is cooperation here: as in most cases, cooperation is needed between the software, that can quickly check what is possible, and the human user who should evaluate which possibilities should be analyzed before giving a final result: after all it is the human that will take the responsibility.

Instead we have listed some antibodies that we had excluded when proceeding manually (anti-C, anti-E, anti-c). Some RBCs which express the target antigens for these antibodies are not reacting, which prompted us to exclude them. Actually we would have taken them into consideration if every other possible explanation had failed, as weak, only partially reacting antibodies. Since a computer can handle huge amounts of data at a time it does not hurt to leave them in from the beginning. In some cases this 'parallel approach' can also be beneficial, examining situations that could have been excluded too quickly.

The scores listed above are the maximum ones possible with the data on panel. A priori we could not exclude the presence of an anti-D reacting at +/-.
At this point the software tries to combine the antibodies and tune the strength required by each to get the best possible fit.

Let us examine a few possible combinations and see how closely they could justify the scores.

Mix	***Comments***
Anti-D	All D+ samples react while all D- samples do not, but sample 1 is reacting (4+) and it should be reacting more weakly that samples 2 and 11 (2+). So this combination is only reasonable if we ignore the scores.
Anti-D+Anti-C	The problem with this combination (as with any combination containing

Mix	*Comments*
	an anti-C specificity) is that RBC sample 5, which is C+, is not reacting. This is not totally unreasonable, because r'r is a weak expression of the C antigen, but somewhat limits the contribute that anti-C reaction can give on the reaction on the first RBC: the anti C must be so weak that it does not react with the r'r RBC, yet so strong that it boosts the reaction on the first RBC from the 2+ that would be caused by the anti-D alone to 4+.
Anti-D+Anti-E or Anti-D+Anti-c	Anti-E and anti-c are possible because their reactions would be covered by anti-D on the RBCs expressing the related antigen strongly (R2R2), yet they should be very weak so that they do not react with RBCs with normal antigenic expressions. Moreover they are not useful to justify the difference between the expected behavior of an anti-D and the observed reactions. So there is no reason to consider these possibilities.
Anti-D+anti-Cw	This seems (and indeed is) a perfect match: C^w antigen is only present on the first RBC were we have too strong a reaction for an anti-D alone. Moreover the reaction due to the anti-C^w alone must be quite strong (3+ or 4+) to justify the jump in reaction from 2+ (anti-D alone) to 4+ (observed).
Anti-D+anti-P1 or Anti-D+anti-Lua	The additional antibody does not help solving the discrepant result on RBC 1, so we discard these possibilities. Eventually we would warn the user for the not excluded anti-Lu^a.
Any combination without anti-D	Such combinations do not even come close to justifying the observed reactions.

Thus we end up with three possibilities in order:

- Anti-D+anti-C^w: which explains everything,
- Anti-D+anti-C: which gets quite close, but requires a weak, only partially reacting anti-C,
- Anti-D alone: which does not explain the differences in the observed score.

Thus Resolvigen suggests the anti-D+anti-C^w antibody mix as the most likely possibility. Resolvigen will list additional possibilities, when there is enough evidence to support them, but here we have a not so well fitting combination with an only partially reacting antibody and a combination that does not fit the observed behavior well. So only the first combination is listed to avoid confusing the user with excessive output.

As we have seen, Resolvigen uses a great deal of expertise to analyze the reactivities: dose effect, variable expression of the antigens, clinical significance of the different antibodies are combined to give a report that is as useful as possible to the user.

Conclusion

Antibody identification is a task that can be completed by a competent technician without any help by a computer; this raises the question: why should we use an antibody identification software.

There are several reason. First of all, no one is borne competent: he becomes so with study and experience; in such cases Resolvigen 3 can act as a tutorial, besides bringing unexpected possibilities to the user attention, it offers on line documentation that can explain why such possibilities were taken into account with a detailed description for over 130 antibodies.

For an expert user a software can complement the deep insight of the user with the breath of analysis and the absolute repeatability of a computer in a manner similar to the way a chess program compares to a chess master: a chess master only analyzes a few moves, usually the more promising, while a chess program analyzes thousands, including the obviously not relevant ones; while the computer lacks the human fantasy and flexibility, it compensate with unfailing attention and breadth of analysis.

Unfortunately the problem of identifying antibodies does not have the clear and simple rules of a chess game, which is why years of experience in the development of antibody identification software are invaluable.