Appendix 2

Supplementary information on analysis of interviews.

All interviews for which audio recordings were made were transcribed verbatim in Microsoft Word. Preliminary codes related to factors that influence land uses were created from three interviews and used as a point of departure to code the remaining interviews during the first round of coding. The codes used were a combination of *a priori* codes generated from background literature and *a posteriori* codes generated from the interview data themselves (Saldaña 2009). The first iteration of coding involved holistic coding, where sets of data were coded as a whole to be divided into more specific codes later, and *in vivo* coding, where direct quotes from the interview transcripts were used as codes (Fig. A2.1). An example of an *in vivo* code created is "no more strength" (used in the context of interviewees choosing not to clear forest for cultivation because of lack of energy to work the land). Codes were then grouped into broader categories derived from the research questions. For example, "no more strength" was put under the category of "reasons for having remaining forest" within the subcategory of "lack of resources". The second round of coding constituted pattern coding, where codes were grouped and renamed as more intangible explanatory codes (Saldaña 2009). For example, "no more strength" and the subcategory of "lack of resources" were then grouped with related codes under the more inclusive category of "availability of resources and labor" as a factor influencing land uses (*i.e.* it was pattern coded as "availability of resources and labor"). Each code was assigned to an inclusive category and new categories were created if there were codes that could not be subsumed under a pre-existing category. This process was repeated to ensure consistency. All coding was carried out on open source TAMS Analyzer software (version 4.45b7ahL).

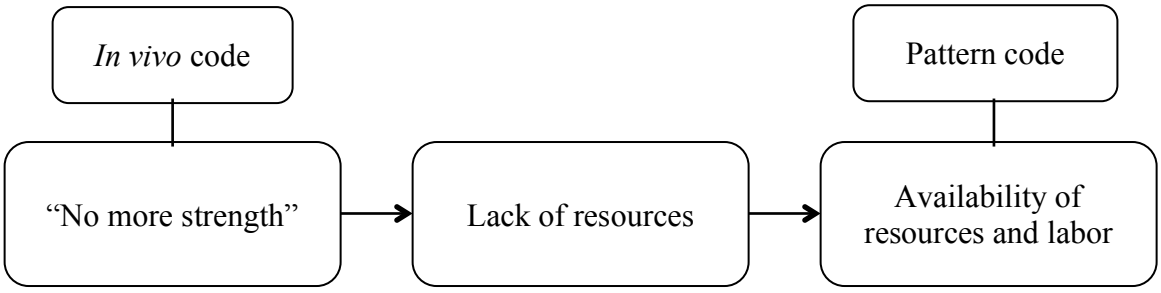

Fig. A2.1. Flow diagram illustrating an example of the process of first- and second-round coding to derive factors that influence the landscape.

Canonical correspondence analyses (CCA; Borcard et al. 2011) were carried out to compare household-level land cover data as a dependent variable from the participatory map to the demographic characteristics of the interviewees, as well as to the social and ecological factors that emerged from interviews as influences on land uses. Three separate CCAs were carried out in RStudio (version 0.98.484): one for demographic characteristics; one for social factors influencing land uses; and one for ecological factors influencing land uses. Canonical correspondence analysis is a weighted redundancy analysis (RDA) that extends regression analysis to multivariate response data by combining multiple linear regression and principal component analysis (PCA) using a matrix of response variables and a matrix of explanatory variables (Borcard et al. 2011). The response matrix of the canonical correspondence analyses in this study consisted of the percent land cover values for each household for which data were available, and the explanatory matrices consisted of demographic characteristics and binary values of social and ecological influences on land uses (Table A2.1).

Only data from those interviewees for whom land cover data were available were used in the CCAs. All the interviewees with parcels to the north of the village had to be removed from the analysis due to the inability of the community members to accurately distinguish between the different parcels in this region. Therefore percent land cover values of these parcels could not be determined. This ambiguity was due to the large number of parcels in a small area and extensive cloud cover preventing referral to the satellite image of the area. Immigrants without land were excluded from analysis and those with land were grouped with landowners. Land inheritors were also grouped with landowners, since they manage their parcels. Youth and women were excluded from the demographic analysis, as they are not the ones who manage the parcels and determine land uses. Any interviewees for whom there were no demographic data for a particular category were removed from the analysis, due to the inability of CCA to manage empty cells. Thus, the interviewees living in Catrigandí, for whom wealth rankings were unavailable, were not included in the demographic CCA. As a result, all landowners with mapped parcels and complete demographic data were included in the demographic CCA (n=10); while all interviewees with mapped parcels were included in the social CCA and in the ecological CCA (n=20).

Stepwise LDAs were not carried out as they did not make a significant difference to correlations.

Table A2.1. Possible inputs used for non-numerical variables in canonical correspondence analysis.

Category	Possible CCA Input

Wealth ranking	1 (poorest) - 4 (wealthiest)
Location	0 (“Site 1”) or 1 (“Site 0”)
Education	0 (none) - 13 (Bachelor’s)
Place of origin	0 (not Darién) or 1 (Darién)
Encroaching frontier settlement, *e.g.*	0 (not mentioned) or 1 (mentioned)

Literature Cited

Borcard, D., F. Gillet, and P. Legendre. 2011. *Numerical Ecology with R*. Springer, New York, New York, USA.

Saldaña, J. 2012. *The coding manual for qualitative researchers*. SAGE Publications, Thousand Oaks, California, USA.