Supplementary Methods

Genetic linkage analysis. Genotype data were checked for incompatibilities using the program PedCheck[1]. LOD scores were computed with the software package FASTLINK[2-4] using the pedigree as a nuclear family with no ancestors or inbreeding (Model 1) and alternatively, assuming that the parents are second cousins (Model 2). The latter approach is called "fuzzy inheritance" in 5, where it was used under similar circumstances of studying families from an isolated population in which parents of sibships with a recessive disease are expected to have a recent common ancestor. Full penetrance was used, except for unaffected individual #6 who had a low neutrophil count, and was coded as having a 95% chance of being unaffected. For the LOD scores shown in **Supplementary Table 2** online, the disease allele frequency was set to 0.001 and marker allele frequencies were set to be all equal. For model 1, the parental genotypes are known, so the marker allele frequencies do not enter into the LOD score calculation. We expected that near the disease gene, affected individuals would be homozygous for the same allele/haplotype, and parents and unaffected siblings would not be homozygous for that allele/haplotype[6]. Therefore, LOD scores are a proxy for detecting segregation that is perfect in this way or close to perfect. The linkage analysis was primarily useful at the early stages to identify regions that were promising, although none of the initial markers was segregating perfectly with the disease. Markers for fine mapping were selected by consulting the Marshfield map[7]. Although the human genome sequence was not complete when the genotyping for this study was undertaken, the particular region of chromosome 1 that proved to contain the gene of interest had been the focus of several other positional cloning studies (e.g., 8 and 9 give useful information about genetic ordering of markers as a byproduct of the hunt that later showed that pycnodysostosis is caused by mutations in cathepsin K[10]). Therefore, we could find many markers for fine mapping and they were reliably ordered.

PCR-amplification of candidate genes. Exon sequences and adjacent 5' and 3' intron sequences of the genes listed in **Supplementary Table 2** online were amplified using specific primer pairs. Standard PCR temperature profile started with 2 min of denaturation at 95 °C followed by 30 cycles of 94 °C for 45 s, 60 °C for 45 s and 72 °C for 60 s, followed by a final elongation step of 72 °C for 6 min PCR products were separated by agarose gel electrophoresis and purified using Qiagen gel extraction kits (Qiagen). For some genes, cDNA was synthesized from total RNA as described below and sequenced on an ABI377 sequencer using specific primers.

DNA Microarray Hybridization and Analysis. Quality and integrity of the total RNA isolated was controlled by running all samples on an Agilent Technologies 2100 Bioanalyzer (Agilent Technologies). For biotin-labeled target synthesis, reactions were performed using standard protocols supplied by the manufacturer (Affymetrix). Briefly, 3 μg total RNA were converted to dsDNA using 100 pmol of a T7T23V primer (Eurogentec) containing a T7 promotor. The cDNA was then used directly in an *in vitro* transcription reaction in the presence of biotinylated nucleotides. The concentration of biotin-labeled cRNA was determined by UV absorbance. In all cases, 12.5 μg of each biotinylated cRNA preparation were fragmented and placed in a hybridization cocktail containing four biotinylated hybridization controls (BioB, BioC, BioD, and Cre) as recommended by the manufacturer. Samples were hybridized to an identical lot of Affymetrix HG_U133A for 16 h After hybridization, the GeneChips were washed, stained with SA-PE and read using an Affymetrix GeneChip fluidic station and scanner.

Microarray Data Analysis. Analysis of microarray data was performed using the Affymetrix GCOS 1.2. The entire data sets are deposited in a MIAME compliant format at Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under series accession number GSE6322. For normalization, all array experiments were scaled to a target intensity of 150, otherwise using the default values of GCOS 1.2. Differential gene expression was calculated using the MAS5

algorithm. Genes were determined as differentially regulated when the change value (measure for a significant expression change) was (I)ncreased for upregulated genes or (D)ecreased for downregulated genes. These criteria must apply for each of the four child/parent comparisons.

Cell lines. EBV B-cell lines were grown in RPMI 1640 (PAA Laboratories) supplemented with 10% heat-inactivated fetal calf serum, 1% penicillin/streptomycin and 10 mM HEPES buffer (all from Gibco / Invitrogen), fibroblasts were grown in DMEM (PAA Laboratories) supplemented with 20% heat-inactivated fetal calf serum and 1% penicillin/streptomycin .

Antibodies and flow cytometry. Peripheral blood mononuclear cells (PBMCs) were stained with the following fluorochrome-conjugated antibodies: CD27-FITC (Dako), anti-IgD-phycoerythrin (Southern Biotechnology Associates), CD19-PC7 (J4.119), CD25-FITC, CD25-PE (both B1.49.9), CD11b-PE (Bear1), CD16-PE (3G8), CD57-FITC (NC1; all from Beckman-Coulter), CD3-FITC (UCHT1), CD27-FITC (M-T271; both from Dako), CD3-APC (S4.1; Caltag Laboratories GmbH), goat-F(ab')2-anti-human-IgD-PE (Biozol), rabbit anti-human-IgM-Cy5 (Dianova). Four color data acquisition was performed with a FACSCalibur (Becton Dickinson). Data analysis (CellQuest, Becton Dickinson) was performed by forward versus side-scatter gating on viable lymphocytes. G-CSF-R-transduced fibroblasts were stained with a PE-conjugated mouse-anti-human G-CSFR antibody (clone LMM741, BD Pharmingen). Data acquisition was performed with a FACScan (Becton Dickinson). Data analysis (CellQuest, Becton Dickinson) was performed by forward versus side-scatter gating on viable fibroblasts.

Functional T-cell studies. T-cells were stimulated with phytohemagglutinin (PHA, 1 µg/ml, Murex Diagnostica), pokeweed mitogen (PWM, 1:800, Biochrom), anti-CD3 mAb HIT3a (10 ng/ml, Becton Dickinson), tetanus toxoid (10 µg/ml, kind gift from K.-D. Hungerer, Behring, Marburg, Germany), tuberculin GT (25 µg/ml, Behring) or alloantigen (mixture of 30 Gy irradiated PBMCs of 6 donors). Cultures were pulsed with 1 Ci ^{3}H-thymidine (0.19 MBq/ml,

Amersham Biosciences) at day 2 and harvested 18 h later. Incorporated radioactivity was determined by means of a Matrix 96 beta counter (Canberra Packard). To assess T-cell cytotoxicity, purified PBMCs were stimulated in RPMI 1640 supplemented with 10% heat-inactivated human AB serum (PAA Laboratories), 10 µg/ml PHA (Sigma-Aldrich) and 50 IU/ml IL-2 (provided by the National Institutes of Health, Bethesda, MD, USA) for 24 h, washed and expanded for 5-7 d in RPMI 1640 supplemented with 10% heat-inactivated human AB serum and 20 IU/ml of IL-2. Effector T-cells were only used when their purity exceeded 95%. For a standard ^{51}chromium release assay, one million L1210 cells (kind gift from R. Jacobs, Hannover Medical School, Hannover, Germany) were labeled for 1 h with 100 µCi ^{51}Cr (Amersham Biosciences), washed twice with PBS and plated at a concentration of 5000 cells/well in triplicates in a 96-well plate. OKT-3-stimulated T-cells were added at various ratios. Maximum release was determined upon incubation with 1% TritonX-100 (Sigma-Aldrich). Supernatants were assessed for ^{51}Cr using a Wallac MicroBeta♦ TriLux detector. To assess granzyme B release, $1x10^5$ cells/well were seeded in a 96-well plate precoated with 100 µl/well of 10 µg/ml OKT-3. Supernatants were harvested and analyzed using a commercially available ELISA kit (Diaclone). Maximum release was obtained upon lysis by 1% TritonX-100. To assess the presence of perforin in T-cells, purified PBMCs were stimulated in RPMI 1640 supplemented with 10% heat-inactivated human AB serum, 10 µg/ml PHA and 50 IU/ml IL-2 for 24 h, washed and expanded for 5 d in RPMI 1640 supplemented with 10% heat-inactivated human AB serum and 20 IU/ml of IL-2. On day 7 5×10^6 cells/well were seeded in a 24-well plate coated with 500 µl/well of 10 µg/ml OKT-3 and incubated for 3-6 h Cells were harvested and lysed in hypotonic buffer containing 20 mM HEPES (pH 7.6), 10 mM KCl, 1 mM $MgCl_2$, 2% glycerin, 0.1% TritonX-100, 0.5 mM DTT (Roche Diagnostics GmbH), 1 mM Pefabloc (Roche), 1 mM sodium-orthovandate (Sigma), 5 g/ml protease inhibitor cocktail (Sigma). Nuclear and cytoplasmic

protein fractions were separated by centrifugation (1800×g). For Western blot analysis, 20 µg of protein were loaded on a 10% SDS gel, separated by electrophoresis and blotted onto nylon membranes. The membranes were exposed to anti-human perforin primary antibodies (Santa Cruz Biotechnology), followed by staining with secondary antibodies conjugated to horse radish peroxidase. The enzymatic reaction was visualized using ECL reagents (ECL Plus kit, Amersham Biosciences).

Electron microscopy. For electron-microscopic studies, 1×10^6 neutrophils were incubated with 3×10^7 *P. aeruginosa* (kind gift from B. Tümmler, Hannover Medical School, Hannover, Germany) for 0, 7.5 and 20 min at 37 °C in RPMI 1640 containing 10% human AB serum and 10 mM HEPES buffer pH 7.4. In addition, primary granulocytes were analyzed directly after hypotonic erythrocyte lysis. For TEM studies $CD8^+$ cells were prestimulated with OKT-3 and incubated with L1210 target cells at a ratio of 30:1. The cells were pelleted at 4 °C at 220×g and first fixed with 2.5% glutaraldehyde (Polysciences) in 0.1 M sodium cacodylate, pH 7.3 (Merck Schuchardt), followed by 2% osmium tetroxide (Polysciences) in the same buffer. After dehydration the specimens were embedded in Epon resin. Thin sections were stained with uranyl acetate and lead citrate. For TEM studies of the skin, thin sections were directly fixed with 2.5% glutaraldehyde in 0.1 M sodium cacodylate, pH 7.3 and processed as described above. Electron microscopy was performed using a Philips electron microscope 301.

RNA Studies. Total RNA was extracted using the NucleoSpin RNA II kit (Macherey & Nagel) or Absolutely RNA^{TM} kit (Stratagene). cDNA was synthesized by mixing 1 g of DNA free total RNA and 50 pmol primer $p(DT)_{15}$, 0.5 mM dNTPs, 5 mM DTT and 1 U Expand Reverse Transcriptase (all from Roche). cDNA was quantified by Real Time PCR in a Light Cycler 2.0 Instrument (Roche) for 30 cycles using the following PCR conditions: 15 s at 95 ^{o}C, 10 s at 60 ^{o}C and 30 s at 95 ^{o}C. p14 cDNA levels were detected by PCR amplification using LightCycler

FastStart DNA Master SYBR Green I Kit (Roche) and either p14 or GAPDH specific primers. For sequencing reactions, cDNA was obtained using the ImProm reverse transcription system (Promega).

Western blot analysis. For Western blot analysis of ERK, pERK, and G-CSF receptor, retrovirus-transduced fibroblasts were grown in 9 cm cell culture dishes to 80% confluency. They were washed, serum-starved in DMEM (PAA) supplemented with 0,5% BSA at 37 °C / 5% CO2 for 24 h, followed by starvation in the absence of BSA. rhG-CSF (Amgen) was added at a final concentration of 50 ng/ml. At the indicated time points, the cells were washed in ice-cold PBS and treated with 100 µl of freshly prepared protein extraction buffer (50 mM HEPES pH 7.4, 1% TritonX-100, 2 mM Na_3VO_4, 100 mM NaF, 1 mM EDTA, 1 mM EGTA, 7% protease inhibitor cocktail (Sigma-Aldrich)) at 4 °C for 1 h.

Mouse anti-human pERK and G-CSF receptor antibodies were purchased from Santa Cruz Biotechnology and rabbit anti-human ERK antibody from Cell Signaling, Danvers, MA, USA. The antibodies were used according to the manufacturers' instructions. To detect p14 in B-cell lines, cells were lysed in buffer (50 mM Tris-HCl, 200 mM NaCl, 1% TritonX-100, 10% Glycerol, 0.5 mM EDTA, 0.5 mM EGTA, 5 mM $Na_4P_2O_7$, 2 mM Na_3OV_4, 50 mM NaF) and cell extracts were subjected to SDS-PAGE and Western blotting using anti-p14 antibodies according to a previously published protocol[14].

Measurement of endosome-nucleus distances in fibroblasts. Data acquisition: Images were acquired using a CCD camera on a Zeiss Axiovert epifluorescence microscope at 16-bit data depth. This was important in order to identify a maximum of endosomes per cell, since cathepsin D positive compartments displayed a high variation in intensities. For the analysis of the endosome-nucleus distances, only GFP positive cells were chosen (visual confirmation). Image preparation: Images containing several complete cells were subdivided in several images with

one cell each, whereas partial cells were not analyzed, but removed from the original images prior to analysis, thus simplifying further analysis of endosome-nucleus distances. Measurement of the endosome-nucleus distances: Endosome-nucleus distances were measured in a semi-automated way using a custom designed Matlab routine. First, nuclei were identified using the DAPI nuclear counterstain. The position of the geometric center of the nucleus was calculated for each cell and used as a parameter representative of the cell center. Individual endosomes were identified on the basis of their fluorescence intensities, using a local thresholding approach, basically identifying regions with local intensity maxima. A global thresholding approach (i.e. intensity thresholding on the whole image) proved insufficient to reliably identify endosomes, since first the local background was substantially different within different areas of a cell, and second, closely spaced endosomes were misidentified as single big endosome. Both sources of mislabeling could largely be eliminated by local thresholding (using the extended maxima transform function of Matlab). It should be noted, however, that although this local intensity-based analysis improved the determination of endosome numbers, it could not be used to determine endosome areas, which were not correctly found by the script. Fortunately, this was not a problem in the current analysis since endosome distribution was used as the only analysis parameter. Visual inspection of several representative masks confirmed the correctness of the endsome counts. The center of each endosome identified that way was determined similarly to the nucleus and the distance to the center of the nucleus was calculated for each endosome using a simple Pythagoras' theorem by overlaying the image on a Cartesian coordinate system. Endosome distances were binned in units of 2 m in the range of 0 to 40 m (from the nuclear center). The number of endosomes in each bin was used for further analysis. A total of 15 GFP control vector transduced healthy donor cells (referred to as HD), 15 GFP control vector transduced patient cells (referred to as #5 non-reconstituted), and 15 p14-GFP transduced patient

cells (referred to as #5 reconstituted) was analyzed and results are given in relative frequencies (%) for each distance (y-axis, see **Figure 6**). statistical analysis: A X^2 independence test was performed on the original data (sum of all endosomes per bin) for HD vs. #5 non-reconstituted, for #5 non-reconstituted vs. #5 reconstituted, and for HD vs. #5 reconstituted resulting in $p < 0.001$ for all three cases, demonstrating the statistical significance of the endosome distributions.

Supplementary References

1. O'Connell, J.R. & Weeks, D.E. PedCheck: A program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.* **63**, 259-266 (1998).
2. Lathrop, G.M. & Lalouel, J.-M. Easy calculations of LOD scores and genetic risks on small computers. *Am. J. Hum. Genet.* **36**, 460-465 (1984).
3. Cottingham, R.W.Jr., Idury, R.M. & Schäffer, A.A. Faster sequential genetic linkage computations. *Am. J. Hum. Genet.* **53**, 252-263 (1993).
4. Schäffer, A.A., Gupta, S.K., Shriram, K. & Cottingham, R.W.Jr. Avoiding recomputation in linkage analysis. *Hum. Hered.* **44**, 225-237 (1994).
5. Hamada, T. *et al.* Lipoid proteinosis maps to 1q21 and is caused by mutations in the extracellular matrix protein 1 gene (*ECM1*). *Hum. Molec. Genet.* **11**, 833-840 (2002).
6. Lander, E.S. & Botstein, D. Homozygosity mapping: A way to map human recessive traits with the DNA of inbred children. *Science* **236**, 1567-1570 (1987).
7. Broman, K.W., Murray, J.C., Sheffield, V.C., White, R.L. & Weber, J.L. Comprehensive human genetic maps: Individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63**, 861-869 (1998).
8. Gelb, B.D., Edelson, J.G. & Desnick, R.J. Linkage of pycnodysostosis to chromosome 1q21 by homozygosity mapping. *Nat. Genet.* **10**, 235-237 (1995).

9. Polymeropoulos, M.H. *et al.* The gene for pycnodysostosis maps to human chromosome 1cen-q21. *Nat. Genet.* **10**, 238-239 (1995).

10. Gelb, B.D., Shi, G.-P., Chapman, H.A. & Desnick, R.J. Pycnodysostosis, a lysosomal disease caused by catehpsin K deficiency. *Science* **273**, 1236-1238 (1996).

11. Hino, M. *et al. Ex vivo* expansion of mature human neutrophils with normal functions from purified peripheral blood $CD34^+$ haematopoietic progenitor cells. *Br. J. Haematol*. **109**, 314-321 (2000).

12. Hamers, M.N., Bot, A.A.M., Weening, R.S., Sips, H.J. & Roos, D. Kinetics and mechanism of the bactericidal action of human neutrophils against *Escherichia coli*. *Blood* **64**, 635-641 (1984).

13. Aarts, L.H.J., Roovers, O., Ward, A.C. & Touw, I.P. Receptor activation and 2 distinct COOH-terminal motifs control G-CSF receptor distribution and internalization kinetics. *Blood* **103**, 571-579 (2004).

14. Teis, D., Wunderlich, W. & Huber, L.A. Localization of the MP1-MAPK scaffold complex to endosomes is mediated by p14 and required for signal transduction. *Dev. Cell.* **3**, 803-814 (2002).

15. Klein, C., Bueeler, H. & Mulligan, R.C. Comparative analysis of genetically modified dendritic cells and cytokine transduced tumor cells as therapeutic cancer vaccines. *J. Exp. Med.* **191**, 1699-1708 (2000).

16. Ory, D.S., Neugeboren, B.A. & Mulligan, R.C. A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc. Natl. Acad. Sci. USA*. **93**, 11400-11406 (1996).