Oral ruxolitinib induces hair regrowth in patients with moderate-to-severe alopecia areata

Julian Mackay-Wiggan,1 Ali Jabbari,1 Nhan Nguyen,1 Jane E. Cerise,1 Charlotte Clark,1 Grace Ulerio,1 Megan Furniss,1 Roger Vaughan,2 Angela M. Christiano,1,3 and Raphael Clynes1

1Department of Dermatology, 2Department of Biostatistics, 3Department of Genetics and Development, Columbia University, New York, New York, USA.

BACKGROUND. Alopecia areata (AA) is a common autoimmune disease with a lifetime risk of 1.7%; there are no FDA-approved treatments for AA. We previously identified a dominant IFN-γ transcriptional signature in cytotoxic T lymphocytes (CTLs) in human and mouse AA skin and showed that treatment with JAK inhibitors induced durable hair regrowth in mice by targeting this pathway. Here, we investigated the use of the oral JAK1/2 inhibitor ruxolitinib in the treatment of patients with moderate-to-severe AA.

METHODS. We initiated an open-label clinical trial of 12 patients with moderate-to-severe AA, using oral ruxolitinib, 20 mg twice per day, for 3–6 months of treatment followed by 3 months follow-up off drug. The primary endpoint was the proportion of subjects with 50% or greater hair regrowth from baseline to end of treatment.

RESULTS. Nine of twelve patients (75%) demonstrated a remarkable response to treatment, with average hair regrowth of 92% at the end of treatment. Safety parameters remained largely within normal limits, and no serious adverse effects were reported. Gene expression profiling revealed treatment-related downregulation of inflammatory markers, including signatures for CTLs and IFN response genes and upregulation of hair-specific markers.

CONCLUSION. In this pilot study, 9 of 12 patients (75%) treated with ruxolitinib showed significant scalp hair regrowth and improvement of AA. Larger randomized controlled trials are needed to further assess the safety and efficacy of ruxolitinib in the treatment of AA.

TRIAL REGISTRATION. Clinicaltrials.gov NCT01950780.

FUNDING. Locks of Love Foundation, the Alopecia Areata Initiative, NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), and the Irving Institute for Clinical and Translational Research/Columbia University Medical Center Clinical and Translational Science Award (CUMC CTSA).


Authorship note: J. Mackay-Wiggan, A. Jabbari, A.M. Christiano, and R. Clynes contributed equally to this work.

Conflict of interest: A.M. Christiano and R. Clynes report personal and other fees from Aclaris Therapeutics, outside of the scope of this work. Additionally, they are inventors for patent US2011/59029 “Methods for Treating Hair Loss Disorders,” licensed to Aclaris Therapeutics.

Submitted: August 3, 2016
Accepted: August 23, 2016
Published: September 22, 2016

Introduction

Alopecia areata (AA) is a major medical problem and is among the most prevalent autoimmune disease in the US, with a lifetime risk of 1.7% (1). AA affects both sexes across all ethnicities and represents the second most common form of human hair loss, second only to androgenetic alopecia (2). AA usually presents with patchy hair loss. One-third of these patients will experience spontaneous remissions within the first year. However, many patients’ disease will progress to alopecia totalis (AT, total scalp hair loss) or alopecia universalis (AU, loss of all body hair). Persistent moderate-to-severe AA causes significant disfigurement and psychological distress in affected individuals (3). In clinical practice, there are no evidence-based treatments for AA (4), yet various treatments are offered, most commonly topical and intralesional steroids, which have limited efficacy.

Our recent mechanistic studies demonstrated a dominant role for type I cellular immunity in AA pathogenesis, mediated by IFN-γ-producing NKG2D-bearing CD8+ cytotoxic T lymphocytes (CTLs) (5). The central role of type I cellular immunity is also reflected in the transcriptional landscape of AA lesional skin in humans and mice, which is dominated by IFN response genes and a CTL signature. These findings provided...
the rationale for therapeutically targeting JAK1/2 kinases in AA, and, indeed, we showed that treatment with JAK inhibitors reversed AA in C3H/HeJ mice and eliminated the type I inflammatory response in the skin (6).

On the basis of our preclinical findings, we initiated a phase II efficacy signal-seeking clinical trial in moderate-to-severe AA, assessing the clinical and immunopathological response to treatment with oral ruxolitinib, a JAK1/2 inhibitor currently FDA approved for the treatment of myeloproliferative disorders.

Results

Efficacy. This study was an open-label clinical trial to investigate ruxolitinib (Jakafi, Incyte Pharmaceuticals), 20 mg orally twice daily, in the treatment of moderate-to-severe AA (Table 1 and Figure 1A). All patients received ruxolitinib for 3–6 months, followed by a 3-month observational phase to assess treatment response durability.

Nine of twelve patients (75%) had significant hair regrowth and achieved the primary outcome of at least 50% regrowth (Table 2). The mean baseline severity of alopecia tool (SALT) score of 65.8% ± 28.0% decreased to a score of 24.8% ± 22.9% at 3 months and to a score of 7.3% ± 13.5% at the end of 6 months of treatment (P < 0.005, Table 2). As a group, the responders exhibited a 92% reduction in hair loss from baseline (Figure 1, B–D; Figure 2; and Supplemental Table 1; supplemental material available online with this article; doi:10.1172/jci.insight.89790DS1), with 7 of the 9 responders achieving over 95% regrowth by end of treatment.

Regrowth was seen in responders as soon as 4 weeks after study medication was initiated and initially presented as variably subtle patchy areas of regrowth, consisting of pigmented terminal hairs, with the exception of one patient (subject 4) with concurrent vitiligo, who exhibited primarily gray hair regrowth.

Table 1. Demographic and treatment time variables overall and by responder status

<table>
<thead>
<tr>
<th></th>
<th>Total (n = 12)</th>
<th>Responder (n = 9)</th>
<th>Nonresponder (n = 3)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean yr [SD])</td>
<td>43.67 (14.41)</td>
<td>44.33 (15.27)</td>
<td>41.67 (14.15)</td>
<td>0.63</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td>58.33 (n = 7)</td>
<td>55.56 (n = 5)</td>
<td>66.67 (n = 2)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>Race (n [%])</td>
<td></td>
<td></td>
<td></td>
<td>0.0091</td>
</tr>
<tr>
<td>White</td>
<td>6 (50)</td>
<td>6 (66.67)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>2 (16.67)</td>
<td>2 (22.22)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>3 (25)</td>
<td>0 (0)</td>
<td>3 (100)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>1 (8.33)</td>
<td>1 (11.11)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Baseline SALT score (mean [SD])</td>
<td>65.63 (26.01)</td>
<td>65.83 (27.96)</td>
<td>65.0 (24.33)</td>
<td>&gt;0.99</td>
</tr>
<tr>
<td>Treatment time (mean wk [SD])</td>
<td>19.67 (4.74)</td>
<td>21.78 (3.07)</td>
<td>13.33 (2.31)</td>
<td>0.0091</td>
</tr>
</tbody>
</table>

Comparisons between baseline and end of treatment were performed using paired Wilcoxon exact sign test for continuous variables or Wilcoxon exact signed-rank test for ordinal variables.

Table 2. Description of variables at baseline and end of treatment, among all subjects, responders only, or nonresponders only

<table>
<thead>
<tr>
<th></th>
<th>Total (n = 12)</th>
<th>Responders (n = 9)</th>
<th>Nonresponders (n = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>End of treatment</td>
<td>Baseline</td>
</tr>
<tr>
<td>SALT score (mean [SD])</td>
<td>65.6 (± 26.0)</td>
<td>21.7 (± 30.3)</td>
<td>65.8 (± 27.96)</td>
</tr>
<tr>
<td>Regrowth (% mean [SD])</td>
<td>0</td>
<td>68.9 (± 42.6)</td>
<td>0</td>
</tr>
<tr>
<td>Visual mean (± SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>671 (± 28.1)</td>
<td>30.1 (± 29.8)</td>
<td>71.7 (± 27.2)</td>
</tr>
</tbody>
</table>

Comparisons between baseline and end of treatment were performed using paired Wilcoxon exact sign test for continuous variables or Wilcoxon exact signed-rank test for ordinal variables.
Of note, the areas of vitiligo in this patient were also noted to improve with ruxolitinib treatment (7). Hair regrowth for all responders increased steadily, with substantial increases each month, resulting in the majority (8 of 9) of responders achieving at least 50% regrowth by the week 12 visit. Responding patients with evidence of regrowth at 3 months continued treatment until they had either achieved 95%–100% regrowth or completed 6 months of treatment.

Durability of responses was assessed in the 3-month follow-up period off treatment. Three of nine responders noted shedding, beginning at week 3 following ruxolitinib discontinuation, and had marked hair loss at week 12 off drug (Supplemental Figure 1); however, hair loss did not reach baseline levels (Figure 1, B and C). Six of nine responders reported increased shedding without major hair loss.

**Biomarker and clinical correlative studies.** Gene expression profiling was performed on skin biopsies

---

**Figure 1. Hair regrowth during and following discontinuation of ruxolitinib treatment.** (A) Patient enrollment flow chart. (B) Severity of alopecia tool (SALT) scores for individual patients during (solid lines) and following cessation of (dashed lines) ruxolitinib treatment. (C) Percentage regrowth for individual patients during and following cessation of ruxolitinib treatment. (D) Predicted (black lines) and actual patient regrowth trajectories (blue lines) from regression models presented in Supplemental Table 1.
taken at baseline and following 12 weeks of treatment, with additional optional biopsies performed earlier in the treatment course. Baseline scalp samples exhibited a distinct gene expression profile when compared with samples taken from unaffected patients (Figure 3A and Supplemental Table 2). Following ruxolitinib treatment, gene expression profiles of AA patient scalp samples clustered more closely with healthy control scalp samples than with baseline AA samples (Figure 3B), indicating global normalization of the AA pathogenic response. Gene expression profiles attributed to the IFN, CTL, and hair keratin (KRT) signatures were assessed in the trivariate AA disease activity index (ALADIN, Figure 3C), a summary index of the AA pathogenic inflammatory response and hair regrowth (8). Importantly, eventual AA responders clustered together on the ALADIN matrix at baseline, sharing high IFN and CTL scores (Figure 3, C and D).

Notably, baseline samples from eventual AA nonresponders exhibited relatively low IFN and CTL scores (Figure 3, D and E) that were not statistically different than normal control samples. Furthermore, in our cohort of AA patients on ruxolitinib treatment, the CTL and IFN signature scores were capable of distinguishing eventual nonresponders and responders at baseline (P < 0.036 and P < 0.036 for CTL and IFN scores, respectively).

Consistent with on-target activity of treatment, skin samples taken following 12 weeks of treatment from responding patients exhibited much lower IFN and CTL scores and clustered much more closely to skin samples taken from normal control patients on the ALADIN matrix (Figure 3D and E). Decreased IFN and CTL scores in biopsies after treatment were demonstrable as early as 2 weeks after the initiation of treatment (Figure 3F).

Adverse events. Ruxolitinib was well tolerated and safely administered in all 12 patients. There were no serious adverse effects, and no patients required discontinuation of therapy. Observed adverse effects were infrequent and included 3 minor bacterial skin infections (in the same patient), 9 episodes of upper respiratory tract infection/allergy symptoms in 7 patients, 1 urinary tract infection, 1 case of mild pneumonia, 1 conjunctival hemorrhage following a surgical procedure, and mild gastrointestinal symptoms. One patient developed lowered hemoglobin, which resolved with dose modification.

Figure 2. Clinical photographs of responder AA patients on ruxolitinib. Pairs of photographs for subjects 1, 2, 3, 4, 8, 9, 10, 11, and 12 are shown as labeled. Photographs labeled “a” in each pair were taken at baseline, and those labeled “b” were taken at the end of treatment with ruxolitinib.
In this proof-of-concept study, 20 mg ruxolitinib twice per day for 3–6 months induced significant hair regrowth in 9 of 12 patients, an overall 75% rate of response to ruxolitinib in the treatment of AA. In contrast, the expected spontaneous remission rates (occurrence of hair regrowth, without treatment) in patients with moderate-to-severe AA is less than 12% based on two randomized controlled trials with similar subject populations (9, 10). Even the most severe forms of alopecia, AT/AU, responded, indicating that the autoimmune process remains pathogenically active and remains reversible with JAK inhibition. Hair regrowth was evident within 1 month in responders and progressed at a rapid rate. Responses were near complete by 6 months of treatment in 8 of 9 responders, suggesting that 6 months of therapy is sufficient to induce maximal clinical remissions in the majority of responders.

In this 9-month study, ruxolitinib was well tolerated. The safety signals in this small study of AA patients, who are otherwise healthy, compare favorably with the prior clinical experience of ruxolitinib in patients with myeloproliferative...
disorders, in which adverse events, particularly those that are hematologically related, are understandably more frequent (11–13), and are consistent with findings from use of tofacitinib in the treatment of patients with psoriasis (14–19). The occurrence of hair shedding in responders following drug cessation suggests that maintenance therapy may be required to sustain remissions. Further clinical evaluation will be required to assess longer durations of therapy and/or investigations of alternative dosing schedules. Larger placebo controlled studies would be required to fully assess the benefit/risk profile of ruxolitinib in AA.

Transcriptional profiling of paired baseline and on-treatment scalp biopsies was both mechanistically and clinically informative. Baseline skin samples from responders had high inflammatory ALADIN IFN and CTL scores, with near normalization after 12 weeks of treatment, indicative of JAK1/2 inhibitor–mediated suppression of the autoreactive CD8 T cell response. Indeed, early ALADIN normalization, as early as week 2 following initiating treatment (Figure 3F), may be predictive of favorable week 12 clinical outcomes. Conversely, nonresponder samples exhibited low baseline IFN/CTL scores and clustered relatively closely to normal patient samples on the ALADIN matrix, suggestive of alternative inflammatory or noninflammatory etiologies of hair loss in these nonresponders (Supplemental Figure 2). One nonresponder had both AA and androgenic alopecia, another nonresponder’s alopecia was consistent with AA histologically but appeared to be a rare diffuse form of the disease, and the final nonresponder exhibited an ophiasis AA pattern. Given our sample size, it is difficult to draw significant conclusions about the efficacy of ruxolitinib in these types of patients. Nevertheless, baseline ALADIN signatures, a measure of type I cellular immunity, may be useful as a predictive biomarker for response to JAK inhibitor treatment, although this needs confirmation in larger well-powered studies, given the caveats described.

Recent single-case reports have described clinical responses in AA patients treated with other JAK inhibitors, including tofacitinib (20), ruxolitinib (21), and baricitinib (22). In addition to this paper and Crispin et al. (23), these proof-of-concept data demonstrate immunopathological reversibility of the type I inflammatory response that underlies AA, even in patients with longstanding or more severe forms of disease, providing a strong rationale for clinical development of oral and/or topical JAK inhibitors for the treatment of AA.

Methods

Study design, oversight, and participants

Study assessments and outcomes. The study’s primary efficacy endpoint was the proportion of responders at end of treatment, defined as those subjects achieving at least 50% regrowth compared with baseline assessed by the SALT score, a standardized, validated method for estimating hair loss in AA (23). Secondary efficacy endpoints included hair regrowth as a continuous variable. Additionally, quality-of-life measures (Dermatology Quality of Life Index and Skindex) were done at regular prespecified intervals but did not show statistical differences in comparisons performed (data not shown). To assess response durability, responders were followed for 3 months after treatment was completed.

Safety analysis was included as a secondary endpoint for all subjects who received at least one dose of ruxolitinib and was monitored as described above at monthly visits.

Biomarker assessment and clinical correlative studies

Biopsies and peripheral blood were obtained at baseline and after 12 weeks for immune monitoring and molecular studies. Several patients provided additional biopsies at intermediate time points during the course of treatment, and one patient provided an additional sample at week 24. Tissues specimens were fixed and stored in PAXgene Tissue Containers (Qiagen). Total RNA was extracted from skin biopsy specimens harvested during the course of the clinical trial using the PAXgene tissue miRNA kit (Qiagen). Library prep was performed for microarray analysis using the Ovation RNA Amplification System V2 and Biotin Encore kits (NuGen Technologies Inc.). Samples were subsequently hybridized to Human Genome U133 Plus 2.0 chips (Affymetrix) and scanned at the Yale Center for Genome Analysis. Library prep and microarray hybridization of RNA extracted from skin biopsies from 3 healthy controls were performed together with the samples from the treated patients for a total of 31 samples. Gene expression analyses included calculation of ALADIN scores, differential expression analysis of the expression levels for the identification of gene expression signatures, principal component analysis, and statistical analysis of the ALADIN scores. Microarray data from the 31 samples have been deposited in GEO under accession number GSE80342.
Statistics overview

Statistical analysis of clinical data. All variables were examined for distributional assumptions and checked for accuracy and for out of range values. Based on our a priori definition, we classified a patient as a responder if the patient experienced 50% or greater hair regrowth from baseline, based on the SALT score at end of treatment. We examined the overall distribution of demographic factors and looked at possible differences between responders and nonresponders, testing for significance using Fisher’s exact test (2 sided) for categorical variables and Mann-Whitney U test for continuous variables. We then examined the change among baseline, end of treatment, and end of study scores for relevant variables overall and for responders and nonresponders employing either a Mann-Whitney U test or Wilcoxon signed-rank test.

To estimate the extent of regrowth across time, we considered both a generalized estimating estimation equation and a mixed model approach to model the repeated-measures data and opted for the latter, given the strong normality assumption of generalized estimating equations and our relatively small sample size. For these mixed models, we first modeled regrowth from baseline to end of treatment, where time (in weeks) was the independent variable, and then, to assess maintenance of the observed effect, modeled regrowth from end of treatment to end of study, where time again was the independent variable. In both models, we specified compound symmetry as the initial covariance structure. For all applicable measures, a P value of less than or equal to 0.05 was considered significant.

Statistical analysis of gene expression data and ALADIN score

In order to examine possible differences among each of the CTL, IFN, and KRT ALADIN scores in responders compared with nonresponders, in responders compared with normal controls, and in nonresponders compared with normal controls, we tested for differences among the 3 groups using a Kruskal-Wallis test, followed by Wilcoxon rank-sum tests, as implemented in the coin package in R. We then examined the change between baseline and ALADIN scores at 12 weeks for responders using Wilcoxon signed-rank test, as implemented in the coin package in R. We further tested for differences in the 3 scores between responders at 12 weeks and normal controls.

ALADIN scores are defined such that mean CTL, IFN, and KRT scores are equal to 0, resulting in mean overall (all patients) scores, responder-only scores, and nonresponder-only scores, corresponding to the mean differences between these and the normal controls. Statistically significant differences were observed between overall scores at baseline and normal controls in CTL (P < 0.0002), IFN (P < 0.005), and KRT (P < 0.0002) scores; between responders-only scores at baseline and normal controls in CTL (P < 0.0004), IFN (P < 0.0004), and KRT (P < 0.0004); between overall scores at week 12 and normal controls in CTL (P < 0.04) and IFN (P < 0.0004); and between responders-only and normal controls in KRT (P < 0.0007) at α = 0.05. No statistically significant difference was observed in IFN scores at baseline overall versus normal controls or in CTL and IFN responders-only scores at week 12 versus normal controls. No statistically significant differences were observed between nonresponders at baseline and normal controls in any of the 3 groups of ALADIN scores.

Changes in ALADIN scores within individual patients were assessed between baseline and week 12. Statistically significant differences were observed between baseline and week 12 overall in the CTL and IFN scores, with KRT scores reaching marginal significance (α = 0.05). CTL scores declined from 8.30 to 1.51 (P < 0.004), IFN scores declined from 31.08 to −0.37 (P < 0.004), and KRT scores increased from −39.36 to −15.02 (P = 0.054) (Table 2). Among responders only, CTL scores declined from 9.37 to 1.6 (P < 0.008), IFN scores declined from 38.37 to 0.24 (P < 0.008), and KRT scores increased from −37.84 to −15.42 (P = 0.039). Statistically significant differences were observed between responders and nonresponders in CTL and IFN scores at baseline (mean score difference = 5.91 and 40.11, P < 0.036 and 0.036, respectively) but not in KRT scores (mean score difference = −19.74, P = 0.22).

Additional methods and statistical analysis can be found in the Supplemental Methods.

Study approval

The study was conceived and conducted by the investigative team at Columbia University. All authors had access to the data and attest to its accuracy and, for the fidelity of this report, to the study protocol. This study was conducted in accordance with Good Clinical Practice, as defined by the International Conference on Harmonization, and in accordance with the ethical principles underlying European Union directive 2001/20/EC and the US Code of Federal Regulations, title 21, part 50 (21CFR50). Prior to study
initiation, the study was reviewed, and approval was obtained from the Columbia University IRB for the protocol and all study-related materials. Freely given written informed consent was obtained from every subject before screening or study-related procedures. Written consent included authorization of the use of photography of patients. Monitoring for regulatory compliance and adherence to the IRB-approved protocol was performed by the Columbia University Clinical Trials Office and the Department of Surgery Regulatory Team. The study was registered on clinicaltrials.gov prior to initiation.

Author contributions
JMW, AJ, AMC, and RC conceived of and designed the study. JMW, NN, CC, GU, and MF evaluated the subjects and/or collected the clinical data. AJ and JEC collected the experimental data. JMW, AJ, NN, JEC, RV, AMC, and RC analyzed the data. JMW, AJ, NN, AMC, and RC wrote the manuscript. All authors reviewed and approved the final version.

Acknowledgments
This work was supported in part by funding from the Locks of Love Foundation, the Alopecia Areata Initiative, the NIH/US Public Health Service (grants from NIAMS, including U01AR067173 and R21AR061881), and the Columbia Skin Disease Research Center (NIAMS P30AR044535). AJ is supported by the Irving Institute for Clinical and Translational Research/CUMC CTSA and NIH/NIAMS (K08AR069110). We are grateful to the AA patients who participated in this study. We thank the National Alopecia Areata Registry and the National Alopecia Areata Foundation for their support of this study. Photographic equipment was provided by Canfield Scientific Inc.

Address correspondence to: Julian Mackay-Wiggan, 161 Fort Washington Avenue, Room 1264, New York, New York 10032, USA. Phone: 212.305.6953; E-mail: jc299@columbia.edu. Or to: Angela Christianso, 1150 St. Nicholas Avenue, Room 307, New York, New York 10032, USA. Phone: 212.851.4850; E-mail: amc65@columbia.edu.


